BMC Health Services Research

RESEARCH

Open Access

A comparative study of neuro-fuzzy and neural network models in predicting length of stay in university hospital

Birgül Yabana Kiremit[1*] and Elif Dikmetaş Yardan[1]

Abstract

Background The time a patient spends in the hospital from admission to discharge is known as the length of stay (LOS). Predicting LOS is crucial for enhancing patient care, managing hospital resources, and optimizing the use of patient beds. Therefore, this study aimed to predict the LOS for patients hospitalized in various clinics using different artificial intelligence (AI) models.

Methods The study analyzed 162,140 hospitalized patients aged 18 and older at various clinics of a university hospital in northern Türkiye from 2012 to 2020. Three soft computing methods—Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Multiple Linear Regression Analysis (MLR)—were employed to estimate LOS using inputs such as medical and imaging services (number of CT, USG, ECG, hemogram tests, medical biochemistry, and number of direct x-rays), demographic, and diagnostic data (patients' age, sex, season of hospitalization, type of hospitalization, diagnosis, and second diagnosis). The LOS predictions utilized single and double-hidden layer ANNs with various training algorithms (Levenberg-Marquardt-LM, Bayesian Regularization-BR and Scaled Conjugate Gradient-SCG) and activation functions (tangent-sigmoid, purelin), ANFIS with Grid Partitioning (ANFIS-GP), and MLR. Model performance was evaluated using the Coefficient of Determination (R^2), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).

Results Of the patients, 54% were male and 43.5% were treated in surgical clinics. The mean age was 55.1 years, with 32.9% of participants aged 65 years or older. Hospital stays were 2–7 days for 39.7% of patients, over 7 days for 30.9%, and 1 day for 29.4%. Neoplasm-related diagnoses (ICD codes) accounted for 25.1% of admissions. Variables influencing LOS were identified through feature selection from patients in various hospital wards. The most significant factors affecting LOS include second diagnosis, the number of hemogram tests, computerized tomography scans (CT), ultrasonography (USG), and direct X-rays. Utilizing these factors, 12 models with varied input variables were developed and analyzed. The double hidden layer ANN model with the Levenberg-Marquardt (LM) training algorithm outperformed the others, achieving R^2 values of 0.854 for training and 0.807 for the test dataset, with RMSE values of 2.397 days and 2.774 days and MAE values of 1.787 days and 1.994 days, respectively. Following ANN-LM, the best results were obtained with ANFIS-GP, while MLR exhibited the lowest performance.

*Correspondence:
Birgül Yabana Kiremit
birgul.yabana@omu.edu.tr

Full list of author information is available at the end of the article

Conclusions Various AI models can effectively predict LOS for patients in different hospital units. Accurate LOS predictions can help health managers allocate resources more equitably across units.

Keywords Artificial neural networks, Adaptive Neuro-Fuzzy Inference System, Multiple linear regression, Length of stay

Introduction

Length of stay refers to the period from admission to the hospital until the patient is discharged [1]. Hospitalizations, or LOS, present complex challenges related to patient health, resource availability, and health policies. Accurate LOS estimation can significantly enhance health system efficiency, reduce costs, optimize hospital service and resource allocation, and ultimately improve patient outcomes [2, 3]. The LOS can be influenced by many factors, such as diagnostic features, patient demographics, and hospital characteristics [4]. Factors influencing LOS include age, gender, marital status, socioeconomic status, insurance coverage, hospital type and size, race/ethnicity, place of residence, and the timing and nature of admission (emergency or urgent). Other considerations are the patient's physical and functional status, discharge status [1], admission details, source, clinical parameters, diagnosis, comorbidities, illness severity, treatments and procedures, medication regimens, and rehabilitation needs [5, 6]. A shorter LOS reduces the risk of infections and adverse drug effects, enhances care quality, and increases hospital profits through better bed management [7]. Conversely, a prolonged LOS raises the risk of nosocomial infections and limits access to inpatient services for other patients. Therefore, estimating the LOS affects the management of the health institution, bed capacity planning, access to health services, and the quality and efficiency of health services [8]. Identifying the most relevant features is crucial in developing effective estimation models for the purpose of predicting the LOS [5]. The enormous amount of medical data available, including patient characteristics, treatment and medication records, and genomic sequences, poses challenges for decision-making [9]. To solve this problem, artificial intelligence can help healthcare providers leverage big data to increase efficiency and achieve results that are difficult to attain using traditional data analyses [10]. Artificial intelligence is a computer system that can perform tasks requiring human intelligence [11]. In modern medicine, artificial intelligence aids clinicians in diagnosis, treatment decision-making, and outcome prediction by analyzing vast information crucial for solving complex clinical problems [12]. Health administrators also benefit from artificial intelligence in health resource management, workforce planning, and hospital bed utilization planning.

This study aims to predict the LOS of patients hospitalized in surgical and medical units of a university hospital between 2012 and 2020 by using AI methods that can reveal linear and non-linear relationships between variables such as ANNs with different training algorithms, ANFIS, and MLR.

The main motivations and contributions of the study are as follows:

- In literature, there are limited studies on LOS prediction with artificial intelligence regardless of the inpatient unit. Our main motivation is to fill this gap in this field.
- This study presents the performances of various training algorithms (LM, BR, SCG) in ANN for predicting LOS, followed by a comparison of ANN results with those from ANFIS and MLR.
- LOS was estimated by using data on the service (procedures) provided to inpatients independent of clinical data.

Artificial intelligence, introduced by John McCarthy [13] in the 1950s and defined as the science and engineering of creating intelligent machines and computer programs, includes commonly used techniques such as machine learning, artificial neural networks, robotics, expert systems, fuzzy logic and natural language processing [14]. Compared to other methods, the use of ANN and ANFIS, which are also discussed in the study, offers some advantages.

The architecture of artificial neural networks is such that information is stored across the entire network. This property ensures that the loss of a few pieces of information in one place does not prevent the network from functioning. ANN exhibits fault tolerance and distributed memory, facilitating the learning of events and the formation of decisions through the analysis of analogous events [15]. While factors and connections that are not clearly visible and ignored can be overlooked by classical methods, artificial neural networks can establish these connections and produce appropriate solutions to the problem. The absence of input data or the capacity for autonomous learning are among the rationales substantiating the superiority of artificial neural networks over classical methods [16]. ANNs are especially suitable for solving non-linear problems. Discriminate analysis, logistic regression and other linear or semi linear techniques typically utilize a limited number of variables that exhibit a significant linear correlation with the dependent variable when constructing their models. ANNs have the capacity to utilize a wider range of information and variables with a very weak linear correlation index [17].

ANFIS, a combination of ANN with fuzzy systems, frequently exhibits the advantage of allowing the final system to be readily translated into a set of if-then rules. Research and applications on the neuro-fuzzy inference strategy have clearly demonstrated the usefulness of neural and fuzzy hybrid systems in areas such as the applicability of existing algorithms for artificial neural networks (ANN) and the direct adaptation of knowledge expressed as a set of fuzzy language rules [18]. ANFIS offers greater flexibility in the membership function and does not necessitate prior human expertise, which is often required in fuzzy systems and may not always be available [19]. Furthermore, the back propagation method employed in neural networks facilitates a significantly reduced merging time by diminishing the size of the search space [20].

Artificial intelligence methods are commonly used to estimate the LOS for patients hospitalized for specific diseases, conditions, or units. Our study estimated the LOS for patients undergoing surgical and medical treatments. However, there are few studies in the literature that predict LOS for all patients across various units. Boff Medeiros et al. conducted a Machine-learning (ML) study that included several pediatric subspecialties [21].

In the existing literature, various factors such as admission data, diagnoses, comorbidities, laboratory results, and particularly the clinical and demographic information of patients have been utilized. In our study, we created a dataset that emphasizes the services provided to patients during their hospitalization, as delays in these services often lead to prolonged hospital stays. Carey et al. examined the factors contributing to extended hospitalization days, finding that 63% of unnecessary days were attributed to non-medical service delays, while 37% were due to medical service delays. Most non-medical service delays (84%) stemmed from challenges in securing a bed in a skilled nursing facility, whereas medical service delays were primarily caused by delayed procedures (54%) and the performance (21%) or interpretation (10%) of diagnostic tests [22]. Consequently, our dataset incorporates service-related data.

Data regarding services, demographic variables, season of admission, diagnoses, and comorbid conditions were collected from 162,140 inpatients at the health institution. After extracting the data, feature selection was conducted, resulting in five features associated with LOS. ANN -with various training algorithms- ANFIS and MLR were used to predict LOS. The performance of the models was evaluated using R^2, RMSE, and MAE. The prediction results from each model were compared and discussed in relation to the existing literature. The findings from this study are anticipated to be beneficial for healthcare managers regarding resource planning, efficient resource management, and enhanced hospital efficiency. Additionally, it is hoped that healthcare professionals will utilize these insights to deliver effective patient treatment and care.

Literature review

Artificial intelligence studies in health

Artificial intelligence has been widely applied in healthcare and medicine [23–25]. In addition, previous studies have focused on predicting patient readmission after a certain period following discharge [26–28], analyzing hospital stays to determine bed capacity and demand [29, 30], and identifying factors that influence LOS [31–33]. Artificial intelligence is frequently employed in LOS estimation, as demonstrated by the studies outlined below.

Artificial intelligence in LOS prediction

Morton et al. estimated the length of stay for 10,000 diabetes patients based on six characteristics: age, gender, race, insurance, admission type, and APR-DRG severity measure. In the study involving Random Forests (RF), Support Vector Machines (SVM), SVM+, Multi-Task Learning (MTL), and MLR. SVM + achieved the highest average ACC (0.68), and AUC (0.76), followed by RF, MTL, and MLR with the next best results [34].

Thompson et al. utilized machine learning to estimate the prolonged LOS in newborns. The dataset comprised information from 17,889 newborns, including details such as date and time of admission, admission number, gender, diagnosis, income level, race, insurance status, and city and state of residence. Various models were employed in the analysis, including Naive Bayes (NB), Logistic Classifier, Multi-layer Perceptron (MLP), Simple Logistic, SVM, Decision Tree (J48), RF, and Random Tree. Among these, the RF model demonstrated the best performance, achieving an ROC value of 0.877 [35].

Maharlou et al. estimated the LOS in intensive care units following cardiac surgery using ANN and ANFIS. The study included 311 heart patients and identified input variables—such as demographic information, medical history, type of surgery, smoking status, and clinical data—based on a literature review. The results indicated that ANFIS provided a more accurate prediction (Mean Squared Error [MSE] = 7 and $R = 0.88$) compared to the ANN (MSE = 21 and $R = 0.60$) [36].

Daghistani et al. estimated the LOS for cardiac patients by analyzing 16,414 patient hospitalizations between 2008 and 2016. Features included were demographics, cardiovascular risk factors, hospitalization and discharge diagnosis, vital signs, and laboratory tests performed during hospitalization. Four different ML methods (ANN, RF, Bayesian Network [BN], and SVM) were utilized for the predictions. The study found that the variables most significantly influencing the prediction of LOS are heart rate at admission, systolic and diastolic blood pressure at admission, age, and insurance status. The RF method yielded the best results (Accuracy: 80%; Sensitivity: 80%; Precision: 80%; F-score: 80%; ROC: 0.94; RMSE: 0.31) [37].

Castineira et al. predicted LOS using machine learning on continuous vital signs and clinical history data from 284 patients in the Pediatric Intensive Care Unit (ICU) at Boston Children's Hospital. They employed a Gradient Boosting Decision Tree (GBDT) method to classify LOS as short- or long-term. The model achieved an AUC of 83% with vital data alone and 90% when combining vital and clinical history data [38].

Kulkarni et al. employed an ANN to predict the LOS and the need for post-acute care in patients with Acute Coronary Syndrome who underwent percutaneous coronary intervention. Different input variables for demographic, clinical, imaging, medication, laboratory, and disease history of the patient were used. This study developed two models and compared their performances. The multi-layer perceptron based model, which predicted long-term hospitalization, achieved 90.87% and 88.36% accuracies in the training and testing datasets, respectively. The model developed for post-acute care achieved 90.22% and 86.31% accuracies in the training and testing datasets, respectively [39].

Wu et al. used four different machine learning methods (RF, BN, SVM, Deep Learning [DL], and GBDT) to predict prolonged LOS in ICU among general intensive care patients. The predictive performance of these methods was compared with the customized Simplified Acute Physiology Score-II (SAPS-II). The data set included patient vital signs, laboratory tests, care plan documents, diagnostic information, and treatment information. The GBDT model demonstrated the best performance in predicting LOS in ICU. The internal and external validation results were as follows: Brier score: 0.164, AUROC: 0.742, AUPRC: 0.537, ECI: 8.224 and Brier score: 0.166, AUROC: 0.747, AUPRC: 0.536, ECI: 8.294 [40].

Fu et al. predicted LOS using characteristics such as demographic, clinical, and second diagnosis of COVID-19 patients. The study used various methods, including MLP, Convolutional Neural Network [CNN], Multi-layer Perceptron with Principal Component Analysis, and Bidirectional Long Short-Term Memory. The CNN method achieved the best results, with an accuracy rate between 73% and 88% [41].

Abd-Elrazek et al. estimated the LOS of patients in the ICU using general medical characteristics collected at admission regardless of patient diagnosis. General demographic characteristics, vital sign characteristics, basic laboratory characteristics, chronic disease characteristics, basic diagnostic characteristics, and Glasgow Coma Scale (GCS) variables were used. As a result of prediction with different ML algorithms, the highest performance was obtained with fuzzy logic with an accuracy rate of 92% [42].

Orooji et al. conducted a study to predict COVID-19 patient LOS using various ANN training algorithms. The study utilized a dataset with 53 features, including demographic information, clinical symptoms, comorbidities, laboratory results, and treatment interventions. They evaluated the performance of 12 different ANN algorithms, analyzing all features as well as those selected through feature selection. The best results were achieved with the BR algorithm, with RMSE values of 1.6213 days and 2.2332 days, respectively [43].

Pan et al. predicted both the prognosis and LOS for patients with traumatic brain injuries. Data from 1,128 patients were utilized, including information on demographic characteristics, in-hospital complications, Glasgow Coma Scale scores, medical history, and whether they were admitted to the intensive care unit. Their study employed Support Vector Regression (SVR), CNN, Backpropagation Neural Networks, ANN, and Logistic Regression models for predictions. SVR had the lowest MAPE value among the models used to calculate the LOS. The MAPE of the SVR model was 9.28%, and the accuracies of the test and external validation datasets were 7.91% [44].

Namavarian et al. compared hospital stay lengths after oral cancer surgery using a statistical model, a machine learning model, and the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) calculator. The dataset included clinical, pathological, and hematological parameters, intraoperative predictors (such as surgical and anesthetic details), and postoperative outcomes. The ML model demonstrated the highest accuracy (validation correlation of 0.48, 4-day accuracy of 70%), outperforming the statistical models: multivariate analysis (0.45, 67%) and least absolute shrinkage and selection operator (0.42, 70%). All models exceeded the ACS-NSQIP calculator's performance (0.23, 59%) [45].

Erdogan Yildirim and Canayaz estimated the LOS in the neonatal intensive care unit. The study dataset includes 453 newborn data and 12 characteristics (level of intensive care, age of the mother, gravida, type of birth, pregnancy week, number of live births, number of abortions, miscarriages or stillbirths, state of multiple pregnancies, baby's weight, baby's gender, and day of admission to intensive care)—a hybrid approach known as Classifier Fusion-LOS was employed. In the initial phase, traditional models such as Logistic Regression, ExtraTrees, RF, K-Nearest Neighbor [KNN], and Support Vector Classifier, along with ensemble models like AdaBoost, Gradient Boosting, XGBoost, and CatBoost, were utilized. RF achieved the highest validation accuracy at 0.94. In the subsequent stage, the Voting Classifier—an ensemble method—was applied, resulting in an increase in model accuracy to 0.96 [46].

The summary of the literature is presented in Table 1.

As depicted in Table 1, studies on predicting LOS reveal a preference for classical models, including RF, SVM, ANN, MLP, NB, MLR, Logistic Regression, and KNN, as well as ensemble models like AdaBoost, Gradient Boosting, XGBoost, and CatBoost. Additionally, datasets featuring patient demographic, clinical, laboratory, and diagnostic information are commonly used for LOS prediction.

This study utilized artificial intelligence methods, specifically ANN, ANFIS, and MLR. Three training algorithms (LM, BR, SCG) were used for ANN predictions, and they were recognized for their quick results, low error risk, and high performance. Few studies estimate LOS using these algorithms together, making this research significant in revealing their predictive performance compared to commonly used techniques. The performance of these training algorithms is compared with ANFIS and MLR results, as no similar studies have been conducted previously. Additionally, the dataset used includes service variables for hospitalized patients, with limited literature addressing this dataset. Most existing studies focus on specific diseases or conditions, while there is a scarcity of models that encompass all age groups and diagnostic spectra in the adult population.

The remainder of the paper is organized into three distinct sections. "Methods" section describes the materials and methods employed to predict LOS. "Results and

Table 1 Comparison between related previous studies in terms of algorithms, features, sample size, best algorithm and prediction results

Study	Algorithms	Features	Sample Size	Best algorithm/ Prediction result
Morton et al. (2014) [34]	RF, SVM, SVM+, MTL, MLR	Demographic characteristics, insurance, admission type, and APR-DRG severity measure.	10,000 Diabetes patients	SVM+ (ACC = 0.68, AUC = 0.76)
Thompson et al. (2018) [35]	NB, MLP, SVM, J48, RF, Log.C, S.Log., RT.	Admission information, socio-demographic characteristics, insurance, diagnosis.	17,889 Newborns	RF (ROC = 0.877)
Maharlou et al. (2018) [36]	YSA, ANFIS	Demographic characteristics, medical history, type of surgery, clinical data and BMI	311 Cardiac surgery patients in intensive care	ANFIS (MSE = 7, R = 0.88)
Daghistani et al. (2019) [37]	ANN, RF, BN, SVM	Demographic characteristics, cardiovascular risk factors, hospitalization and discharge diagnosis, vital signs, and laboratory tests	16,414 Cardiac patients	RF (Sensitivity = 0.80, Accuracy = 0.80, AUROC = 0.94)
Wu et al. (2021) [40]	RF, BN, SVM, DL, GBDT	Vital sign measurements, laboratory tests, care plan documentation, diagnosis information, treatment information, and others.	200,859 ICU patients	GBDT (In internal validation Brier score = 0.164), AUROC = 0.742; In external validation Brier score = 0.166), AUROC, 0.747)
Fu et al. (2021) [41]	MLP, CNN, PCA + MLP, BiLSTM	Demographic characteristics, clinical data, and second diagnosis	100,000 COVID-19 patients	CNN (Accuracy = 88,49%, MSE = 0,27)
Abd-Elrazek et al. (2021) [42]	CART, ANN, TB, RF, FL, SVM, KNN, NB	Demographic characteristics, vital signs, laboratory tests, Symptoms, chronic disease, current complaint and diagnosis, Glasgow Coma Scale	233 ICU patients	FL (Accuracy = 92%, Specificity = 92%)
Orooji et al. (2022) [43]	12 Different ANN training algorithms-LM, BR, BFGS, RP, SCG, CGB, CGF, CGP, OSS, GDX, GDM, GD	Demographic characteristics, clinical symptoms, comorbidities, laboratory results, treatment interventions.	343 COVID-19 patients	BR (For whole features RMSE = 1.6 days; For selected features RMSE = 2.2 days)
Pan et al. (2023) [44]	SVR, CNN, BNN, ANN, LR.	Demographic characteristics, in-hospital complications, Glasgow Coma Scale scores, medical history, whether	1,128 Traumatic brain injuries patients	SVR (MAPE = 9.28%, Accuracy = 7.91%)
Namavarian et al. (2024) [45]	MVA, LASSO, ML, ACS-NSQIP	Clinical, pathological, and hematological parameters, intraoperative predictors, and postoperative outcome	837 Oral cavity cancer patients	ML (Correlation = 0.48, 4-day Accuracy = 70%)
Erdogan Yildirim and Canayaz (2024) [46]	LR., ET, RF, KNN, SVC, AdaBoost, Gradient Boosting, XGBoost, CatBoost	Level of intensive care, information about mother and birth, baby's weight, gender, and day of admission to intensive care	453 Newborn in neonatal intensive care	RF (Accuracy = 0.94, F1-Score = 0.90)

Discussion" section summarizes the results obtained from ANN, ANFIS, and MLR, followed by a discussion of these findings. Finally, "Conclusion" section presents the overall conclusions of the paper.

Methods

Data

The study data were collected at a university hospital in Samsun, Türkiye. This study included patients 18 years and older who were hospitalized between 2012 and 2020. A total of 162,140 inpatient records were obtained from the hospital's automation system. The data, which were in Excel format, were combined by merging the rows for each day that the patient stayed in the hospital. Age, gender, diagnosis, secondary diagnosis, type of hospitalization, the season of hospitalization, number of procedures (number of CT, number of USG, number of ECG, number of hemogram tests, number of Medical Biochemistry, number of Direct X-Ray) data of inpatients were obtained from the health institution for use in the study (Table 2).

The first feature is a numerical variable representing the number of days a patient spends in the hospital, which will serve as the output variable. The variables that include age and gender are demographic, while the next two features indicate the season of service and whether the care provided was medical or surgical. The following two features pertain to the patient's diagnosis group (classified according to ICD-10) and whether there is a secondary diagnosis accompanying the primary diagnosis. The last six attributes relate to the resources provided by the hospital for inpatient services and are therefore considered administrative attributes.

Before feature selection, we assessed the data set for missing or incorrect entries. It was found that some birth dates used to calculate patients' ages were either incorrect or unspecified, leading to the exclusion of those patients. Conversely, patients who had not undergone specific inpatient procedures (e.g., CT, USG, hemogram test, D.X-Ray) were included in the study. Naemi et al. noted that patients with missing values might provide valuable information that can enhance the performance of prediction models [47]. Thus, we included patients who had not undergone one procedure while having received others.

Correlation analysis was used to determine the relationship between data from the health institution and LOS. Feature selection identified the following variables for the study: the number of CT scans, ultrasound (USG), hemograms, direct X-rays, and secondary diagnoses among inpatient procedures. The dataset includes five attributes, comprised of both numeric and categorical variables.

Method

Artificial intelligence techniques, including ANN, the hybrid artificial intelligence method (ANFIS), and the MLR, were used in this study.

Artificial Neural Networks (ANN)

ANN is a model that can accurately process complex processes [13] and allow prediction and generalization by processing incomplete and non-linear data [14]. An ANN method consists of an input layer, one or more hidden layers, and an output layer. The input layer contains information from the environment, while the hidden and output layers consist of neurons that perform linear and non-linear computations, respectively [48].

Prior to training, artificial neural networks (ANNs) require the setting of hyperparameters, which can be numerical or functional [49]. This study employs the number of layers, activation function, and neurons per layer as hyperparameters. ANNs are structured as either

Table 2 List of variable used in this study

Attribute Name	Description	Variable Type	Values
Length of Stay (LOS)	The length of stay of patients in hospital.	Numerical	$1 \leq LOS \leq 109$
Age	The patient's age at admission	Numerical	$18 \leq Age \leq 106$
Gender	The patient gender (Female or Male)	Categorical	0 or 1
Season	The season at the patient's admission	Categorical	0–1–2 or 3
Type of stay	The surgerical or medical treatment patient	Categorical	0 or 1
Diagnosis	The chapter regarding the ICD of diagnosis of the patient on admission	Categorical	Between 0–24
Second Diagnosis	It indicates whether the patient has a secondary diagnosis.	Categorical	0 or 1
Number of CT (CT)	The total CT number of the patient.	Numerical	$0 \leq CT \leq 22$
Number of USG (USG)	The total number of ultrasonography for the patient.	Numerical	$0 \leq USG \leq 14$
Number of ECG	The total number of electrocardiogram for the patient.	Numerical	$0 \leq ECG \leq 30$
Number of Hemogram tests (Hem.)	The patient's hemogram test number.	Numerical	$0 \leq Hem \leq 126$
Number of Medical Biochemistry tests (Med. Biochem.)	The patient's medical biochemistry test number.	Numerical	$0 \leq$ Med. Biochem ≤ 114
Number of Direct X-Ray (D.X-Ray)	The total number of X-rays of the patient.	Numerical	$0 \leq$ D.X-Ray ≤ 48

feedforward or feedback networks, depending on neuron connectivity. Feedforward networks feature unidirectional information flow from input to output [50], while feedback networks exhibit dynamic behavior, with neuron outputs influenced by previous outputs [51]. Feedforward networks are frequently favored due to their structural flexibility, representational power, and the availability of numerous learning algorithms [52]. The network architecture used in this study is feedforward.

Artificial neural network training typically involves partitioning the dataset into training and test sets, or training, test, and validation sets. The training set allows the network to learn input-output relationships. Test and validation sets, smaller than the training set [53], evaluate the network's performance on unseen data. In this study, the dataset was divided into 70% for training and 30% for testing, analyzed across 100–500 iterations. The number of iterations was determined using a trial-and-error approach, stopping when the error value stabilized.

An artificial neural network is a computational intelligence model capable of performing both classification and regression tasks. It utilizes self-learning optimization to predict outputs based on given inputs [54]. ANNs are particularly effective in estimation analyses [55, 56]. Within artificial neural networks, various regularization techniques are employed alongside the Backpropagation training algorithm to minimize small errors. These techniques help ensure smoother operation of the ANN and reduce the likelihood of overfitting [57]. LM and BR are two regularization methods used to address the overfitting issue in ANNs; these algorithms can achieve lower mean squared errors compared to others [58]. Additionally, the Scaled Conjugate Gradient (SCG) algorithm calculates the derivative of the error function with respect to each weight, adjusting the weights to minimize the error.

The LM algorithm is recognized as a standard technique for solving non-linear least squares problems. It consists of a combination of gradient descent or steepest descent and the Gauss-Newton method [59]. The LM algorithm is a high-order adaptive algorithm and determines the best direction in which to move the weights to reduce the mean square error [60].

Bayesian Regularization is a successful algorithm for revealing potentially complex relationships and can be used to build a robust model in quantitative studies [61]. BR has an objective function that includes the residual sum of squares and the sum of squared weights to minimize prediction errors and obtain a better generalized model [62, 63].

Scaled Conjugate Gradient utilizes second-order information from a feed-forward neural network such as the LM algorithm. This algorithm requires more iterations than other conjugate gradient algorithms, but the number of calculations in each iteration is significantly reduced because no row search is performed. This mechanism makes the algorithm faster than other algorithms [64].

Activation functions, which are expressed as a mathematical function that enables the transfer of information from one neuron to another [65], are effective in the performance of artificial neural networks. Tansig, logsig, and purelin activation functions are frequently used in ANN [66]. In this study, tangent-sigmoid (tansig) is used in the hidden layer, and linear function (purelin) is used in the output layer.

Information on the construction of the study models is presented in Table 3.

This study used ANN with a single input layer, one or two hidden layers, and a single output layer to predict LOS (Table 3). Although it has been suggested that a single hidden layer may be adequate for accurate model estimation [67], in this study, two hidden layer models were also considered to compare their results. Trial and error determined the number of neurons in the hidden layers. Various training algorithms are available for neural network training [68]. In this case, the Neural Network was trained using three different algorithms: LM, BR, and

Table 3 Structure of the models

Number of input	1	2	3	4
Number of Hidden Layer	1 and 2			
Number of neurons in hidden layer	Between 4 and 11			
Training algorithm	1 Levenberg-Marquardts algorithm (trainlm)			
	2 Scaled Conjugate Gradient algorithm (trainscg)			
	3 Bayesian Regularization algorithm (trainbr)			
Activation function	Sigmoid (tansig) in the hidden layer, Linear function (purelin) in the output layer			
Model Structures	1–4–1	2–4–1	3–4–1	4-4–1
	1–6–1	2–6–1	3–6–1	4–6–1
	1–8–1	2–8–1	3–8–1	4–8–1
	1–10–1	2–10–1	3–10–1	4–10–1
	1–5-7-1	2–5-7-1	3–5-7-1	4–5-7-1
	1–9-11-1	2–9–11-1	3–9-11-1	4–9-11-1

SCG. The input and output layers used tangent-sigmoid (tansig) and linear (purelin) activation functions, respectively. The number of hidden layer nodes ranged from 4 to 11 to obtain the optimal training network. The number of neurons in the hidden layer was determined by trial and error [69]. The data were normalized between 0 and 1 to eliminate unit problems. The data were randomly divided into the training dataset (70%) and the testing dataset (30%). The model hyperparameters were determined using the training dataset, and their prediction performance was evaluated using the test dataset.

The algorithms to be used in the training of ANN were preferred for the study since they eliminate the overfitting problem, give a lower error, are fast, and have high accuracy in terms of performance. As far as we know, there is no literature study comparing the performance results of these three training algorithms for the duration of patient hospitalization.

The Adaptive Neuro-Fuzzy Inference System (ANFIS)

Jang [20] introduced the ANFIS, a hybrid AI technique combining fuzzy logic and ANN [70]. ANFIS integrates the decision-making mechanism of fuzzy logic with the learning ability and relational structure of the ANNs.

ANFIS applies the Takagi Sugeno inference model, which consists of IF-THEN rules, to establish a mapping between inputs and outputs [71]. The method enhances estimation performance by updating the parameters utilized in the fuzzy rules and leveraging the learning capabilities of the neural network.

ANFIS comprises five layers [72]: fuzzy membership functions, fuzzification multiplication, normalization, aggregation, and output function layers. This architecture includes fixed and adaptive nodes; membership function (Layer 1) and output (Layer 4) nodes are adaptive, while others are fixed. Layer 1 nodes represent membership functions (e.g., triangle, trapezoid, Gaussian) [73]. Layer 2, the rule layer, outputs the firing level of each rule [72]. Layer 3 uses fixed nodes to normalize firing strengths [74]. Layer 4 represents the conclusion of each fuzzy rule [73]. Layer 5 sums incoming signals to produce the overall model output [74].

Various methods can be employed to significantly reduce runtime and enhance the performance and interpretability of the ANFIS method. Subtractive Clustering (SC), Grid Partitioning (GP), and Fuzzy C-Means Clustering (FCM) are commonly used techniques in the ANFIS. The ANFIS used the GP method in this study. Grid partitioning involves dividing an area into a grid-like structure to prevent regions overlapping in the input space. This approach is particularly suitable when dealing with a small number of input variables or when input space dimensions are limited. To give a simple example, in a model with 10 input variables, when each input variable is divided into 2 member functions, 2^{10} = 1024 specific areas are formed, and a rule is created for each specific area. The total number of rules is 1024, which is a very complex structure. Therefore, GP is used in cases where the number of input variables is small [75]. In this study, models with 1, 2, 3, and 4 inputs were created (2^1=2; 2^2=4; 2^3=8 and 2^4=16 rules), and this prevented the complex structure that would occur in the use of the ANFIS-GP method. The data calculation process in this study incorporates triangular, trapezoidal, and Gaussian membership functions.

Multiple Linear Regression (MLR)

Regression analysis is a statistical technique to estimate the relationships between different variables [76]. MLR can simultaneously model the relationships between multiple dependent variables and numerous independent variables [77].

The equation for an MLR is given by [78].

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta p X p + \varepsilon \quad (1)$$

where

Y = dependent variable;

X_1, X_2,...X_p = explanatory variables

β_1, β_2... β_p = the slope coefficients for each explanatory variable;

β_0 = constant term;

ε = the model's error term.

Data preparation (Preprocessing)

The dataset, which comprises 13 different features of hospitalized patients, was processed prior to its use in estimating LOS. This dataset was obtained from the health institution in separate Excel files, which were then combined and checked for missing patient data. It was found that some patients had missing diagnosis codes and incorrectly entered birth dates, leading to their exclusion from the study. Although not every patient utilized all the services included in the dataset, and some examination and imaging services were not provided at all, these patients were still included in the study. This decision was based on the understanding that other available data from patients with missing values could influence the performance of the prediction model developed for LOS [47]. Therefore, these patients were included in the study.

The dataset includes variables such as gender, type of hospitalization, diagnosis, having a second diagnosis, and season of hospitalization. These variables are categorical. ANN can analyze numerical data, so these categorical data were converted to numerical data in Microsoft Excel. The dataset contains a total of 13 features, 5 of which are categorical and 8 are numerical. Feature selection was performed to determine the features to be used in LOS estimation from these 13 features of the patients.

Factors such as input combinations, model structures, basic parameters, and performance criteria affect the performance of ANN models. Therefore, it is very important to define the input variables correctly when creating a correct prediction model. Selecting inappropriate input variables can significantly diminish the model's effectiveness. Statistical methods were employed to identify the most important variables influencing LOS, the dependent variable. In literature, feature selection can involve artificial intelligence-based methods, expert opinions, and statistical techniques. In our study, we conducted Pearson correlation analysis, which was favored by Gul and Guneri, Dogu et al., and Orooji et al. for feature selection, to determine the input variables [43, 79, 80]. Based on the results of this analysis, variables that exhibited a strong relationship with LOS were chosen as input variables for the study. Following feature selection, five features related to LOS were identified: the number of CT scans, direct X-rays, USG, hemogram tests, and the presence of a secondary diagnosis.

In order to apply the prediction algorithms on the dataset with 5 input and 1 output variable and to prevent overfitting, 70% of the data were randomly divided to be used in the training ($n = 113498$) and 30% in the testing ($n = 48642$) phase. After the data was divided into training and testing data, prediction models were created to compare the performances of the ANN, ANFIS, and MLR methods. Three activation functions were selected to be used in the dataset training with ANN: LM, BR, and SCG. It was found that BR and LM had better performance than traditional methods in terms of both speed and overfitting problems [61], and SCG was preferred for the predictions to be made because it gives fast results by avoiding time-consuming row search in each iteration of the calculation. Before training the created models, the models were scaled to the range of 0–1, and each algorithm was applied to the training data set. The trained models were validated on the test data set. Different

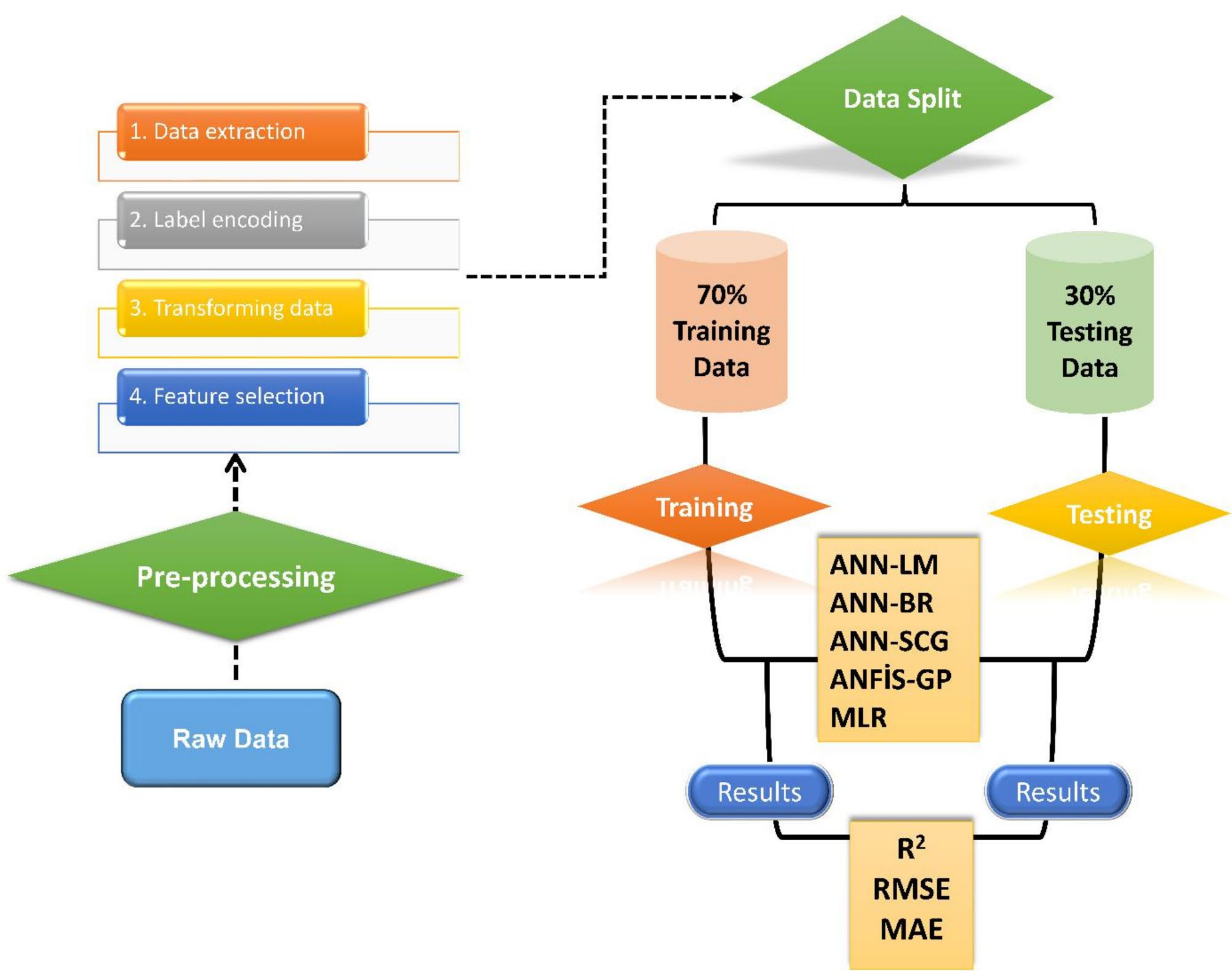

Fig. 1 Flowchart of the LOS estimation model

indicators, R^2, MAE, and RSME, were used to evaluate the performance of the models. The performances of the models were compared, and the one with the best results was identified (Fig. 1).

Modelling with ANN, ANFIS and MLR methods

A training-test analysis strategy was used in this study to create models for the ANN, ANFIS, and MLR methods. These models were based on variables that were determined to impact the LOS in the previous stage. Various input combinations create prediction models to achieve the best prediction.

A feed-forward MLP with an input layer, one or two hidden layers, and an output layer was used to make predictions using the artificial neural network. The network was trained using LM, BR, and SCG training algorithms. Different single and double-layer network structures and iterations are employed. The transfer functions used in the hidden and output layers were sigmoid (tansig) and linear (purelin) functions. The analysis of the ANFIS models involved the use of the GP training algorithm and Trimf, Trapmf, and Gaussmf membership functions. In all the models, 3*3 and 4*4 membership functions were used. MATLAB 2018 software was used to

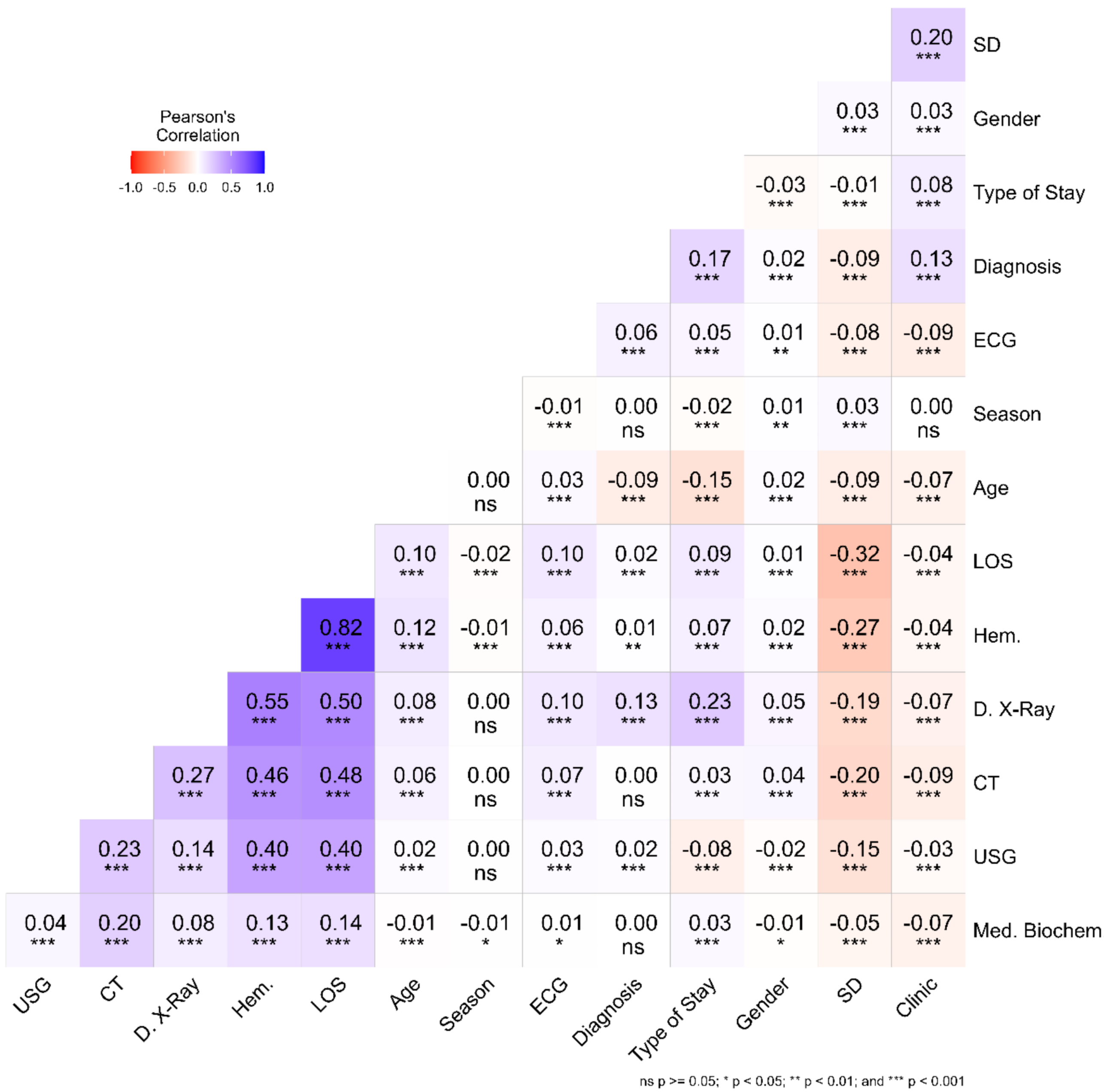

Fig. 2 Correlation between input variables

Table 4 Descriptive statistics for the dataset used in the LOS prediction models

		Min.	Max.	Mean	SE.	SD.	Kurt.	Skew.
Training Dataset	Length of Stay	1	109	6.918	0.024	8.21	12.85	2.79
	Number of CT	0	22	0.60	0.004	1.42	16.33	3.42
	Number of Direct X-Ray	0	48	1.21	0.007	2.52	32.24	4.43
	Number of Hemogram tests	0	126	3.93	0.020	6.86	22.20	3.74
	Number of USG	0	14	0.16	0.001	0.50	45.68	4.96
	Second Diagnosis*							
Testing Dataset	Length of Stay	1	108	6.93	0.04	8.32	14.95	2.97
	Number of CT	0	21	0.61	0.01	1.46	17.56	3.51
	Number of Direct X-Ray	0	46	1.21	0.01	2.52	33.41	4.50
	Number of Hemogram tests	0	106	3.94	0.03	6.95	24.65	3.93
	Number of USG	0	10	0.17	0.00	0.52	33.43	4.63
	Second Diagnosis*							

*Categorical Data, *SE* Standard Error, *SD* Standard Deviation

predict the LOS. The number of neurons in the network was determined through a trial-and-error process. The performance of the network is evaluated after training and testing. Before training, the models used in this study were scaled to the range of 0–1. The scaling equation given in Eq. 2 was used for this purpose.

$$X_{norm} = \frac{X_a - X_{\min}}{X_{\max} - X_{\min}} \tag{2}$$

X_{norm} is a variable's normalized value, X_a is a variable's measured value, and $X_{\max}$ and $X_{\min}$ are the measured maximum and minimum values of a variable, respectively.

Evaluation of model performance

The LOS was estimated using the modeling methods described above. Various techniques have been employed to assess the estimate accuracy and compare models. In this study, statistical performance evaluation criteria such as the R^2, RMSE, and MAE were calculated and used to compare the results. The equations for these performance criteria are as follows (Eqs. 3–5) [81]

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Z_m - Z_p)^2}{\sum_{i=1}^{n} (Z_m - Z_a)^2} \tag{3}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (Z_p - Z_m)^2}{n}} \tag{4}$$

$$MAE = \frac{\sum_{i=1}^{n} (|Z_p - Z_m|)}{n} \tag{5}$$

Where;

Z_m = the measured value;

Z_p = the estimated value;

Z_a = the average value measured;

n = the number of data.

The sociodemographic and descriptive analyses and the MLR analysis were conducted using IBM SPSS Statistics for Windows 25.0 (IBM Corp., Armonk, NY, USA) statistical software. ANN and ANFIS analyses were performed using Matlab R2018a. Additionally, correlation analysis was conducted using RStudio 4.1. software and heatmap plots were created using Google Colaboratory.

Results and discussion

This study utilized data from 162,140 patients hospitalized at Ondokuz Mayıs University Health Application and Research Centre Adult Hospital between 2012 and 2020. Of these patients, 46% were female and 54% were male. Among the hospitalized patients, 43.53% were

Table 5 Input combinations to be used in the Estimation of the LOS with different estimation methods

Models	Inputs
M1	Hemogram
M2	Hemogram, Computed Tomography
M3	Hemogram, Second Diagnosis
M4	Hemogram, Ultrasonography
M5	Computed Tomography, Direct X-ray, Ultrasonography
M6	Hemogram, Second Diagnosis, Ultrasonography
M7	Hemogram, Second Diagnosis, Computed Tomography
M8	Hemogram, Computed Tomography, Ultrasonography
M9	Hemogram, Direct X-ray, Second Diagnosis, Ultrasonography
M10	Second Diagnosis, Direct X-ray, Computed Tomography, Ultrasonography
M11	Hemogram, Direct X-ray, Computed Tomography, Ultrasonography
M12	Hemogram, Second Diagnosis, Computed Tomography, Direct X-ray

admitted to the surgical departments, while 56.74% were admitted to the internal medical departments. In the surgical departments, 73.9% of the patients were aged 18–64, and 26.1% were 65 years and older. For patients admitted to internal medical departments, the percentages for the age groups 18–64 years and 65 years and older were 61.8% and 38.2%, respectively. The mean age of the admitted patients was 55.1 years.

Regarding LOS, 39.7% of the patients were hospitalized for 2–7 days, 30.9% for more than 7 days, and 29.4% for 1 day. The highest number of hospitalizations occurred in the autumn and winter seasons (26.1% and 25.3% respectively). The distribution of inpatient admissions based on ICD-10 diagnosis groups revealed that most patients (25.1%) were admitted with neoplasm-related diagnosis codes. Furthermore, 60.1% of the inpatients received secondary diagnoses.

Pearson correlation analysis was conducted to establish the relationship between variables to construct models for use in ANN, ANFIS, and MLR (Fig. 2). Correlation analysis revealed that certain variables, such as the number of CT scans, direct x-rays, hemogram tests, USG, and the presence of a secondary diagnosis, exhibited significant positive or negative correlations with the LOS. However, these variables were excluded from the study due to their weak association with patient age, gender, admission season, principal diagnosis, type of hospitalization, number of medical biochemistry, and number of ECGs. The variables that displayed relationships among themselves were randomly divided into two datasets: training

Table 6 Performance results of the LOS prediction models with ANN

Model	Training Algorithms	Network Structure	Training Dataset			Testing Dataset		
			R^2	RMSE	MAE	R^2	RMSE	MAE
M1	LM	1–9–11-1	0.729	3.990	2.472	0.671	5.736	3.482
	SCG	1–9–11-1	0.729	3.996	2.479	0.672	5.735	3.487
	BR	1–9–11-1	0.729	3.992	2.473	0.672	5.733	3.481
M2	LM	2–9–11-1	0.826	2.869	2.073	0.821	2.881	2.080
	SCG	2–9–11-1	0.819	2.923	2.134	0.815	2.931	2.142
	BR	2–9–11-1	0.825	2.873	2.078	0.822	2.879	2.080
M3	LM	2–9–11-1	0.863	2.344	1.742	0.803	2.955	2.025
	SCG	2–9–11-1	0.861	2.359	1.765	0.802	2.964	2.044
	BR	2–5–7-1	0.863	2.345	1.742	0.803	2.956	2.025
M4	LM	2–9–11-1	0.716	4.350	2.745	0.673	4.841	2.974
	SCG	2–9–11-1	0.715	4.353	2.748	0.673	4.838	2.977
	BR	2–9–11-1	0.710	4.391	2.794	0.677	4.818	3.007
M5	LM	3–9–11-1	0.696	2.171	1.160	0.626	3.496	2.535
	SCG	3–6–1	0.687	2.657	2.142	0.627	3.507	2.574
	BR	3–5–7-1	0.695	2.621	2.095	0.633	3.477	2.530
M6	LM	3–9–11-1	0.854	2.397	1.787	0.807	2.774	1.994
	SCG	3–9–11-1	0.849	2.435	1.839	0.803	2.807	2.046
	BR	3–9–11-1	0.854	2.399	1.789	0.806	2.776	1.998
M7	LM	3–9–11-1	0.724	4.429	2.780	0.684	4.377	2.767
	SCG	3–9–11-1	0.716	4.487	2.848	0.680	4.408	2.826
	BR	3–9–11-1	0.723	4.438	2.793	0.686	4.362	2.776
M8	LM	3–9–11-1	0.727	4.264	2.674	0.628	5.176	2.896
	SCG	3–6–1	0.718	4.334	2.774	0.690	4.737	2.950
	BR	3–9–11-1	0.726	4.271	2.677	0.696	4.690	2.879
M9	LM	4–9–11-1	0.871	2.272	1.682	0.814	2.882	1.963
	SCG	4–6–1	0.863	2.348	1.784	0.804	2.946	2.061
	BR	4–5–7-1	0.871	2.277	1.687	0.813	2.884	1.966
M10	LM	4–9–11-1	0.728	2.597	2.032	0.694	2.852	2.169
	SCG	4-4–1	0.720	2.637	2.074	0.694	2.854	2.197
	BR	4–9–11-1	0.728	2.599	2.037	0.698	2.831	2.169
M11	LM	4–9–11-1	0.776	3.001	1.880	0.623	7.156	4.494
	SCG	4–5–7-1	0.760	3.105	1.982	0.669	6.923	4.561
	BR	4–9–11-1	0.775	3.005	1.882	0.672	6.868	4.471
M12	LM	4–9–11-1	0.885	2.320	1.554	0.697	5.647	3.475
	SCG	4–8–1	0.871	2.451	1.691	0.693	5.712	3.533
	BR	4–9–11-1	0.884	2.296	1.515	0.698	5.641	3.472

Table 7 Performance results of the LOS prediction models with ANFIS-GP

Inputs	Models	Training Algorithms	Training Dataset			Testing Dataset		
			R^2	RMSE	MAE	R^2	RMSE	MAE
		Trimf (3)	0.727	4.010	2.478	0.673	5.736	3.483
		Trapmf (3)	0.718	4.076	2.643	0.663	5.855	3.621
Hem.	M1	Gaussmf (3)	0.722	4.043	2.590	0.667	5.812	3.570
		Trimf (4)	0.728	3.998	2.477	0.672	5.733	3.483
		Trapmf (4)	0.720	4.059	2.622	0.665	5.823	3.593
		Gaussmf (4)	0.723	4.037	2.577	0.668	5.787	3.549
Hem.-CT	M2	Trimf (3-3)	0.824	2.885	2.089	0.820	2.894	2.093
		Trapmf (3-3)	0.814	2.962	2.226	0.809	2.978	2.234
		Gaussmf (3-3)	0.817	2.939	2.185	0.812	2.953	2.193
		Trimf (4-4)	0.824	2.883	2.088	0.820	2.895	2.093
		Trapmf (4-4)	0.816	2.948	2.202	0.811	2.962	2.210
		Gaussmf (4-4)	0.818	2.927	2.164	0.814	2.939	2.170
Hem.-SD	M3	Trimf (3-3)	0.863	2.349	1.746	0.802	2.957	2.028
		Trapmf (3-3)	0.857	2.397	1.834	0.793	3.023	2.124
		Gaussmf (3-3)	0.857	2.393	1.825	0.795	3.012	2.112
		Trimf (4-4)	0.863	2.348	1.746	0.802	2.959	2.029
		Trapmf (4-4)	0.857	2.394	1.830	0.794	3.019	2.120
		Gaussmf (4-4)	0.858	2.388	1.817	0.795	3.008	2.104
Hem.-USG	M4	Trimf (3-3)	0.713	4.368	2.754	0.677	4.817	2.965
		Trapmf (3-3)	0.704	4.440	2.918	0.672	4.860	3.099
		Gaussmf (3-3)	0.709	4.402	2.856	0.671	4.860	3.058
		Trimf (4-4)	0.714	4.362	2.756	0.674	4.832	2.973
		Trapmf (4-4)	0.707	4.417	2.887	0.665	4.907	3.090
		Gaussmf (4-4)	0.710	4.389	2.831	0.296	8.740	3.098
CT-D.X-Ray-USG	M5	Trimf (3-3-3)	0.694	2.626	2.099	0.528	3.891	2.549
		Trapmf (3-3-3)	0.687	2.656	2.137	0.623	3.527	2.569
		Gaussmf (3-3-3)	0.690	2.644	2.127	0.614	3.549	2.567
		Trimf (4-4-4)	0.695	2.623	2.097	0.477	4.147	2.560
		Trapmf (4-4-4)	0.689	2.648	2.131	0.624	3.523	2.570
		Gaussmf (4-4-4)	0.692	2.636	2.117	0.344	5.166	2.596
Hem.-SD-USG	M6	Trimf (3-3-3)	0.853	2.403	1.792	0.807	2.778	1.998
		Trapmf (3-3-3)	0.846	2.461	1.880	0.798	2.841	2.089
		Gaussmf (3-3-3)	0.848	2.447	1.859	0.800	2.827	2.066
		Trimf (4-4-4)	0.848	2.448	1.864	0.789	2.900	2.075
		Trapmf (4-4-4)	0.847	2.456	1.874	0.798	2.837	2.082
		Gaussmf (4-4-4)	0.849	2.440	1.850	0.796	2.855	2.060
Hem.-SD-CT	M7	Trimf (3-3-3)	0.722	4.443	2.786	0.686	4.366	2.768
		Trapmf (3-3-3)	0.716	4.490	2.874	0.677	4.423	2.858
		Gaussmf (3-3-3)	0.719	4.463	2.838	0.682	4.394	2.821
		Trimf (4-4-4)	0.722	4.443	2.802	0.684	4.377	2.786
		Trapmf (4-4-4)	0.718	4.474	2.855	0.679	4.415	2.838
		Gaussmf (4-4-4)	0.720	4.457	2.829	0.681	4.395	2.812
Hem.-CT-USG	M8	Trimf (3-3-3)	0.711	4.388	2.844	0.679	4.822	2.994
		Trapmf (3-3-3)	0.448	6.060	4.453	0.434	6.372	4.500
		Gaussmf (3-3-3)	0.656	4.786	3.231	0.604	5.339	3.316
		Trimf (4-4-4)	0.712	4.374	2.831	0.677	4.836	2.982
		Trapmf (4-4-4)	0.547	5.487	3.962	0.500	5.998	4.045
		Gaussmf (4-4-4)	0.671	4.679	3.146	0.382	7.262	3.273
Hem.-D.X-Ray-SD-USG	M9	Trimf (3–33-3)	0.870	2.281	1.691	0.691	3.682	1.985
		Trapmf (3–33-3)	0.865	2.332	1.777	0.796	3.006	2.067
		Gaussmf (3–33-3)	0.866	2.316	1.752	0.745	3.318	2.046

Table 7 (continued)

Inputs	Models	Training Algorithms	Training Dataset			Testing Dataset		
			R^2	RMSE	MAE	R^2	RMSE	MAE
SD- D.X-Ray -CT-USG	M10	Trimf (3–33-3)	0.727	2.604	2.037	0.677	4.815	2.200
		Trapmf (3-3-3-3)	0.721	6.459	4.881	0.604	6.717	4.983
		Gaussmf (3-3-3-3)	0.724	6.463	4.881	0.631	6.724	4.987
Hem.- D.X-Ray -CT-USG	M11	Trimf (3-3-3-3)	0.763	3.085	1.981	0.667	6.950	4.557
		Trapmf (3-3-3-3)	0.506	4.459	3.379	0.434	8.972	6.148
		Gaussmf (3-3-3-3)	0.732	3.286	2.225	0.601	7.451	4.821
Hem.-SD-CT- D.X-Ray	M12	Trimf (3-3-3-3)	0.873	1.017	1.684	0.694	1.076	3.537
		Trapmf (3-3-3-3)	0.587	1.090	3.250	0.495	1.178	6.085
		Gaussmf (3-3-3-3)	0.835	1.036	1.982	0.664	1.175	6.477

(70%) and test (30%) datasets for ANN, ANFIS, and MLR. Table 4 presents the descriptive statistics, including the minimum (Min.), maximum (Max.), mean, standard error (SE), standard deviation (SD), kurtosis (Kurt.), and skewness (Skew.) coefficients, for both the training and test datasets.

As shown in Table 4, the LOS values for both the training and test datasets ranged from 1/day to 109/days and 1/day to 108/days, respectively. The CT numbers vary between 0 and 22 pieces and 0–21 pieces, while the Direct X-ray numbers range from 0 to 48 pieces to 0–46 pieces. Additionally, the number of hemogram tests varies between 0 and 126 and 0–106. Lastly, the number of ultrasonography performed ranged from 0 to 14 pieces to 0–10 pieces, respectively. It is worth noting that the models used for the analyses were created using data divided into training and test datasets (Table 5).

Various input combinations were considered to assess the influence of each factor on the LOS. With this aim, the input combinations utilized during the training and testing phases are presented in Table 5. According to that, 12 models were created with input parameters ranging from 1 to 4. The dataset for these models was normalized between 0 and 1, and their performance was assessed. The artificial neural network made LOS predictions using three training algorithms: LM, SCG, and BR. The results of the prediction analysis are shown in Table 6.

The training and testing datasets performance results of the ANN with different training algorithms that estimated LOS are presented in Table 6. The second column lists the training algorithms used, and the third column specifies the optimum number of hidden layer neurons. The table shows that the ANN-LM (3–9–11-1) model, which used three input parameters, achieved the highest accuracy in the testing period according to the R^2, RMSE, and MAE criteria. The three inputs in this model are the number of hemogram tests, the second diagnosis, and the number of ultrasound tests. The double hidden layer model consisted of 9 neurons in the first layer, 11 neurons in the second layer, and one output. The accuracy criteria of the best models during the testing period were $R^2=0.807$, RMSE = 2.774 days, and MAE = 1.994 days. BR and SCG were the other training algorithms that produced the best statistical results. Although multiple hidden layer models are used in ANN, studies have stated that a single hidden layer model is sufficient [82, 83]. However, they stated that the MLP is the best model for solving complex problems because it overcomes the disadvantage of a single-layer perceptron by adding more hidden layers [84]. The LM algorithm achieved the best results in our study, followed by the BR algorithm. In the study conducted by Orooji et al., which estimated the LOS of COVID-19 patients, various ANN training algorithms were compared, with the best results achieved using BR following LM [43].

The ANFIS models were developed using the GP training algorithm, incorporating various membership functions (MFs) like triangular, trapezoidal, and Gaussian to predict LOS. The performance of the ANFIS-GP across different models in both training and test datasets, along with the optimal number of MFs and parameter values, is presented in Table 7.

The results of the LOS prediction with ANFIS-GP are shown in Table 7. The ANFIS (3, trimf) model, which employs three triangle membership functions for the inputs (number of hemogram tests, second diagnosis, and number of ultrasonography), demonstrates superior performance compared to the other ANFIS models during the testing phase. The accuracy metrics for the best model of M6 in the testing phase were $R^2=0.807$, RMSE = 2.778 days, and MAE = 1.998 days.

The multiple regression equations developed based on various independent variables are presented in Table 8, and the accuracy of the models for the training and testing datasets is shown in Table 9. Among the MLR models used to predict LOS in the test dataset, M6 had the lowest RMSE and MAE (2.845 days and 2.105 days, respectively) and the highest coefficient of determination (0.797).

The performance results for the three modeling approaches, namely ANN-LM, ANFIS-GP, and MLR, are comprehensively summarized in Table 10. Upon

Table 8 Mathematical equations from MLR models for estimation of LOS

Model	Inputs	Mathematical Equations	R^2	MAE	RMSE
M1	Hem.	LOS = 2.530 + 0.953×HEM. *SH = [0.014]***[0.002]****	0.657	3.561	5.658
M2	Hem.-CT	LOS = 2.434 + 0.666×CT + 0.907×HEM *SE = [0.011]**[0.008]**[0.002]***	0.807	2.270	2.994
M3	Hem.-SD	LOS = 2.730 + 0.937×HEM − 1.006 × SD *SE = [0.012]***[0.001]***[0.016]****	0.794	2.131	2.980
M4	Hem.-USG	LOS = 2.808 + 0.940×HEM + 1.287×USG *SE = [0.015]***[0.002]***[0.029]****	0.667	3.156	4.868
M5	CT-D.X-Ray-USG	LOS = 2.823 + 1.738×CT + 1.246×D.X-Ray + 4.068×USG *SE = [0.012]**[0.009]**[0.005]**[0.025]***	0.631	2.568	3.429
M6	Hem.-SD-USG	LOS = 3.175–1.249 × SD + 0.915×HEM + 0.224×USG *SE = [0.014]**[0.019]**[0.002]**[0.022]***	0.797	2.105	2.845
M7	Hem.-SD-CT	LOS = 3.564 + 0.901×HEM + 0.687×CT −1.482 × SD *SE = [0.021]***[0.002]***[0.011]***[0.029]****	0.671	2.910	4.462
M8	Hem.-CT-USG	LOS = 2.666 + 0.878×HEM + 0.659×CT + 1.167×USG *SE = [0.015]***[0.002]***[0.010]***[0.029]****	0.681	3.069	4.774
M9	Hem.-D.X-Ray-SD-USG	LOS = 3.074–1.237 × SD + 0.233× D.X-Ray + 0.878×HEM + 1.096×USG *SE=[0.015]**[0.020]**[0.005]**[0.002]**[0.021]***	0.805	2.121	2.939
M10	SD- D.X-Ray -CT-USG	LOS = 3.770–2.067 × SD + 1.699×CT + 1.178× D.X-Ray + 4.022×USG *SE = [0.015]**[0.019]**[0.009]**[0.005]**[0.023]**	0.692	2.192	2.861
M11	Hem.- D.X-Ray -CT-USG	LOS = 2.069 + 0.785×HEM + 0.642×CT + 1.100×USG + 0.155× D.X-Ray *SE=[0.011]***[0.002]**[0.008]***[0.022]***[0.005]****	0.666	4.289	6.413
M12	Hem.-SD-CT- D.X-Ray	LOS = 2.768 + 0.890×HEM + 0.622×CT − 1.032 × SD + 0.179× D.X-Ray *SE=[0.012]***[0.002]**[0.007]***[0.016]***[0.004]****	0.696	3.517	5.493

*: $p < 0.05$; **: $p < 0.01$

Table 9 Evaluating the performance of MLR models

Inputs	Model	Training Dataset			Testing Dataset		
		R^2	MAE	RMSE	R^2	MAE	RMSE
Hem.	M1	0.713	2.894	4.169	0.657	3.561	5.658
Hem.-CT	M2	0.811	2.264	2.984	0.807	2.270	2.994
Hem.-SD	M3	0.856	1.870	2.411	0.794	2.131	2.980
Hem.-USG	M4	0.696	3.026	4.500	0.667	3.156	4.868
CT-D.X-Ray-USG	M5	0.682	2.186	2.693	0.631	2.568	3.429
Hem.-SD-USG	M6	0.845	1.899	2.471	0.797	2.105	2.845
Hem.-SD-CT	M7	0.710	2.924	4.537	0.671	2.910	4.462
Hem.-CT-USG	M8	0.706	2.957	4.424	0.681	3.069	4.774
Hem.-D.X-Ray-SD-USG	M9	0.862	1.827	2.360	0.805	2.121	2.939
SD- D.X-Ray -CT-USG	M10	0.708	2.106	2.691	0.692	2.192	2.861
Hem.- D.X-Ray -CT-USG	M11	0.757	2.366	3.303	0.666	4.289	6.413
Hem.-SD-CT- D.X-Ray	M12	0.863	1.916	2.609	0.696	3.517	5.493

Table 10 Comparison of the best prediction methods

Method	Inputs	Model Parameters	Training			Testing		
			R^2	RMSE	MAE	R^2	RMSE	MAE
ANN-LM	Hem.-SD-USG	LM (3–9–11-1)	0.854	2.397	1.787	0.807	2.774	1.994
ANFIS-GP	Hem.-SD-USG	Trimf (3-3–3)	0.853	2.403	1.792	0.807	2.778	1.998
MLR	Hem.-SD-USG		0.845	2.471	1.899	0.797	2.845	2.105

reviewing the outcomes of these models, it is evident that the ANN-LM model outperformed the others, attaining the highest R^2 value of 0.807, which indicates a strong correlation between the predicted and actual values. Additionally, it also recorded the lowest RMSE of 2.774 days, suggesting a higher degree of accuracy in its predictions, alongside the lowest MAE of 1.994 days, further emphasizing its effective performance. In contrast, the

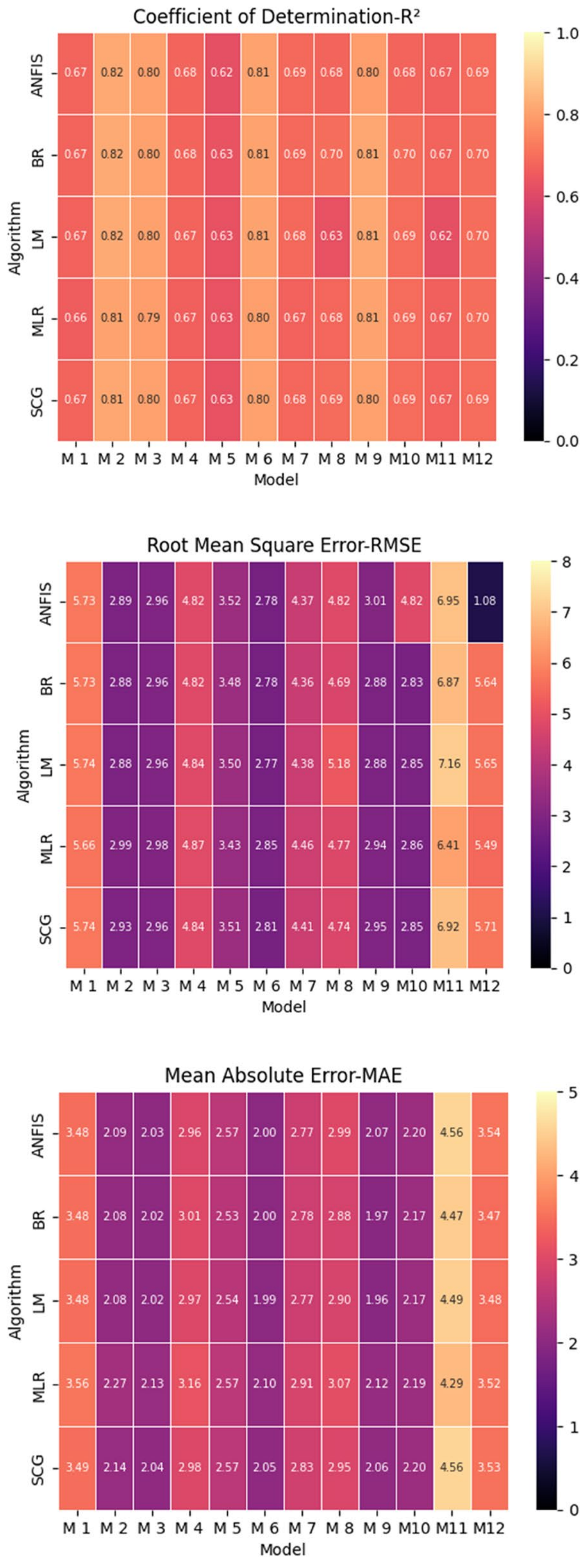

Fig. 3 Performance metrics (R^2, RMSE, and MAE) of the best algorithms for different input combinations in the test dataset

MLR model showed less impressive results, recording the lowest R^2 value of 0.797, which implies a weaker relationship between the predicted and actual outcomes. It also faced the disadvantage of having the highest RMSE of 2.845 days and the highest MAE of 2.105 days, indicating reduced accuracy compared to its counterparts. These findings suggest that MLR struggles to model the non-linear relationships inherent in the data adequately, a limitation commonly associated with linear models in scenarios where variables exhibit complex interactions. Overall, these results validate the use of ANN-LM as a powerful tool for predictive modeling, particularly in applications requiring high accuracy and precision. The ranking of model performance — MLR as the least accurate, followed by ANFIS-GP and ANN-LM as the most accurate — reinforces the importance of selecting appropriate modeling techniques tailored to the complexity of the dataset. Future studies may explore the integration of additional features or hybrid models to enhance predictive capabilities further and address the limitations observed in models like MLR and ANFIS-GP.

Maharlou et al. employed ANFIS and ANN to estimate the LOS for patients in the ICU following cardiac surgery, finding that ANFIS outperformed ANN [36]. In contrast, our study observed the opposite result, possibly attributable to differences in sample size in the data set used.

Figure 3 shows the R^2, RMSE, and MAE values of the best-performing models for different input combinations in the test dataset. The model with the combination of Hem, SD, and USG inputs exhibited the highest R^2 and lowest RMSE and MAE in the testing dataset (Fig. 3). Additionally, using the hemogram input variable in each of the developed models increased the accuracy of the models. This underscores the hemogram's strong predictive power and its potential role as a key determinant in the target outcome. The enhanced accuracy achieved by incorporating Hemogram may be attributed to its biological or physiological relevance to the predicted variable, thereby improving the model's ability to capture essential patterns and reduce errors.

Among the three prediction methods (ANN, ANFIS, and MLR) utilized in this study, the ANN yielded the most favorable outcomes. This finding is consistent with other studies that estimate LOS using artificial intelligence methods [85–89], thus confirming the superior predictive ability of artificial neural networks.

Figure 4 shows a scatter plot of the ANN-LM algorithm, where M6 performed best with Hem.-SD-USG input variables. Most data points cluster near the 1:1 line.

In our study, we estimated patient LOS across various disease groups using artificial intelligence methods, achieving acceptable results ($R^2=0.807$, MAE = 1.99 days). A related study focused on pediatric patients in different sub-specialties, also estimating LOS for various

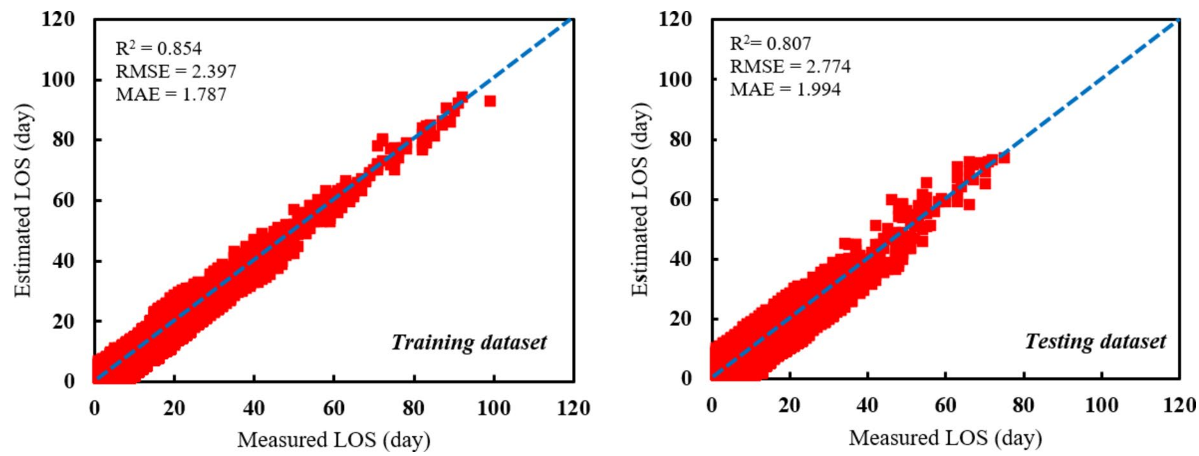

Fig. 4 Scatter plots between measured and estimated data for LOS

diseases. While the performance indicators were adequate, the MAE was slightly elevated, attributed to the complexity of estimating LOS for diverse conditions [21].

Conclusion

This study used AI (ANN, ANFIS, MLR) to predict patient LOS in clinics. ANN-LM, using hemogram tests, secondary diagnoses, and ultrasonography as inputs, was the most accurate (1.9 days error), outperforming MLR based on R^2, RMSE, and MAE. Hemogram tests significantly improved model performance. These findings help healthcare administrators optimize resource allocation and service delivery. This study demonstrates AI's feasibility for accurate LOS prediction. This work establishes AI's value for LOS prediction, enabling better resource management and advancing healthcare analytics.

The limitations of the present study are as follows:

This study, limited by its specific input variables and single-institution data source, cannot be generalized to other healthcare institutions. Furthermore, factors influencing LOS are subject to change over time.

Future research

Future research should be conducted with a larger dataset that includes more clinical, economic, and sociodemographic data. Other artificial intelligence models should be explored, models should be applied in real time, integrated with EHRs, and cost-benefit analyses should be performed. Multi-institutional studies in different settings will increase generalizability. Political and environmental changes may affect changes in patient LOS; research should be extended accordingly.

Abbreviations

AI	Artificial intelligence
ACS-NSQIP	American College of Surgeons National Surgical Quality Improvement Program
ANFIS	Adaptive Neuro-Fuzzy Inference Systems
ANFIS-GP	ANFIS with Grid Partitioning
ANN	Artificial Neural Networks
BiLSTM	Bidirectional Long Short-Term Memory
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BN	Bayesian Network
BNN	Backpropagation Neural Networks
BR	Bayesian Regularization
CART	Classification And Regression Trees
CGB	Conjugate Gradient with Powell/ Beale
CGF	Conjugate Gradient Fletcher-Powell
CGP	Conjugate Gradient with Polak-Ribiére
CNN	Convolutional Neural Network
CT	Computer Tomography
DL	Deep Learning
DT	Decision Tree
ECG	Electrocardiogram
ET	ExtraTrees
FCM	Fuzzy C-Means Clustering
FL	Fuzzy Logic
GBDT	Gradient Boosting
GCS	Glasgow Coma Scale
GDX	Gradient Descent Variable Learning Rate
GDM	Gradient Descent with Momentum
GD	Gradient Descent
GP	Grid Partitioning
ICU	Intensive care units
KNN	K-Nearest Neighbor
Kurt	Kurtosis
LM	Levenberg-Marquardt
LOS	Length of Stay
Log.C	Logistic Classifier
LR	Logistic Regression
MAE	Mean Absolute Error
MVA	Multivariate Analysis
ML	Machine Learning
MLP	Multi-Layer Perceptron
MLR	Multiple Linear Regression Analysis
MTL	Multi-Task Learning
MSE	Mean Square Error
NB	Naïve Bayes
OSS	One Step Secant
QN	Quasi-Newton
PCA + MLP	Multi-layer Perceptron with Principal Component Analysis
RF	Random Forest
RB	Resilient Backpropagation
RT	Random Tree
R^2	Coefficient of Determination
RMSE	Root Mean Square Error
S.Log	Simple Logistic
SAPS	Simplified Acute Physiology Score
SC	Subtractive Clustering
SCG	Scaled Conjugate Gradient
SD	Standard deviation

SE	Standard Error
Skew	Skewness
SVC	Support Vector Classifier
SVM	Support Vector Machine
SVR	Support Vector Regression
TB	Tree Bagger
USG	Ultrasonography

Acknowledgements
Not applicable.

Authors' contributions
BYK contributed to data collection, analyzed the data, prepared the figures and tables, and wrote the original draft of the manuscript, while EDY designed and supervised the project. All authors discussed the results and commented on the manuscript.

Funding
This work was supported by the Scientific Research Projects Unit of Ondokuz Mayis University. Project number: PYO.SBF.1904.21.003.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
In order to conduct the study, the necessary Ethical Committee approval was obtained from the Ondokuz Mayıs University Clinical Research Ethics Committee in 2020 (No: OMU KAEK 2020/613). The data used in the study consists of administrative data routinely collected by the hospital and obtained from the hospital database. There was no need to collect additional data from the patients. The study is based on the examination of retrospective medical records, and the patient data were obtained from the hospital at the time the patients were discharged. The study data were anonymized in a way that ensures patient confidentiality (Türkiye Personal Data Protection Law - Article 28 (1)-b) [90]. Personal data that could compromise individual security and privacy were not included in the dataset. Therefore, individual informed consent was not requested. Information about informed consent was provided to the Ondokuz Mayıs University Clinical Research Ethics Committee. Permission was obtained from the Ondokuz Mayıs University Health Practice and Research Center for obtaining the study data.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Healthcare Management, Faculty of Health Sciences, Ondokuz Mayis University, Atakum, Samsun 55200, Türkiye

Received: 10 July 2024 / Accepted: 20 March 2025
Published online: 27 March 2025

References

1. Ravangard R, Arab M, Zeraati H, Rashidian A, Akbarisari A, Mostaan F. Patients' length of stay in women hospital and its associated clinical and non-clinical factors, Tehran, Iran. Iran Red Crescent Med J. 2011;13(5):309.
2. Ayyoubzadeh SM, Ghazisaeedi M, Rostam Niakan Kalhori S, Hassaniazad M, Baniasadi T, Maghooli K, Kahnouji K. A study of factors related to patients' length of stay using data mining techniques in a general hospital in Southern Iran. Health Inform Sci Syst. 2020. https://doi.org/10.1007/s13755-020-0099-8.
3. Almeida G, Brito Correia F, Borges AR, Bernardino J. Hospital length-of-stay prediction using machine learning algorithms—a literature review. Appl Sci. 2024;14(22):10523.
4. Tsai PFJ, Chen PC, Chen YY, Song HY, Lin HM, Lin FM, Huang QP. Length of hospital stay prediction at the admission stage for cardiology patients using artificial neural network. J Healthc Eng. 2016. https://doi.org/10.1155/2016/7035463.
5. Symum H, Zayas-Castro JL. Prediction of chronic disease-related ınpatient prolonged length of stay using machine learning algorithms. Healthc Inf Res. 2020;26:20–33.
6. Hempel L, Sadeghi S, Kirsten T. Prediction of ıntensive care unit length of stay in the mımıc-ıv dataset. Appl Sci. 2023;13:6930.
7. Baek H, Cho M, Kim S, Hwang H, Song M, Yoo S. Analysis of length of hospital stay using electronic health records: A statistical and data mining approach. PLoS ONE. 2018. https://doi.org/10.1371/journal.pone.0195901.
8. Pollack MM, Holubkov R, Reeder R, Dean JM, Meert KL, Berg RA,… and, Tamburro R. PICU length of stay: factors associated with bed utilization and development of a benchmarking model. Pediatr Crit Care Med. 2018. https://doi.org/10.1097/PCC.0000000000001425.
9. Saadatmand S, Salimifard, Mohammadi R, Kuiper A, Marzban M, Farhadi A. Using machine learning in prediction of ICU admission, mortality, and length of stay in the early stage of admission of COVID-19 patients. Ann Oper Res. 2023. https://doi.org/10.1007/s10479-022-04984-x.
10. Beam AL, Drazen JM, Kohane IS, Leong TY, Manrai AK, Rubin EJ. Artificial intelligence in medicine. N Engl J Med. 2023. https://doi.org/10.1056/NEJMe2206291.
11. Jiang Y, Li X, Luo H, Yin S, Kaynak O. Quo vadis artificial intelligence? Discover Artif Intell. 2022. https://doi.org/10.1007/s44163-022-00022-8.
12. Ramesh AN, Kambhampati C, Monson JR, Drew PJ. Artificial intelligence in medicine. Ann R Coll Surg Engl. 2004;86(5):334–8.
13. McCarthy J. What is artificial intelligence. 2007. https://hub.cnetworks.info/wp-content/uploads/2023/07/whatisai.pdf. Access January 2023.
14. Omitaomu OA, Niu H. Artificial intelligence techniques in smart grid: A survey. Smart Cities. 2021. https://doi.org/10.3390/smartcities4020029.
15. Mijwel MM. Artificial neural networks advantages and disadvantages. Mesopotamian J Big Data. 2021;2021:29–31.
16. Güven I. Demand forecasting with artificial intelligence methods in retail apparel industry. (Master Thesis). Karabük: Karabük University; 2020.
17. Grossi E, Buscema M. Introduction to artificial neural networks. Eur J Gastroenterol Hepatol. 2007;19(12):1046–54. https://doi.org/10.1097/MEG.0b013e3282f198a0.
18. Walia N, Singh H, Sharma A, ANFIS. Adaptive neuro-fuzzy inference system-a survey. Int J Comput Appl. 2015;123(13):32–38.
19. Basri M. Medical image classification and symptoms detection using neuro fuzzy, in electrical-mechatronics and automatic control.Doctoral dissertation, Universiti Teknologi Malaysia, Faculty of Electrical Engineering; 2008.
20. Jang JSR. ANFIS: adaptive-network-based fuzzy inference system, in IEEE transactions on systems, man, and cybernetics. 1993;23(3):665–85. https://doi.org/10.1109/21.256541.
21. Boff Medeiros N, Fogliatto FS, Karla Rocha M, Tortorella GL. Predicting the length-of-stay of pediatric patients using machine learning algorithms. Int J Prod Res. 2023;63(2):483–496.
22. Carey MR, Sheth H, Braithwaite RS. A prospective study of reasons for prolonged hospitalizations on a general medicine teaching service. J Gen Intern Med. 2005;20:108–15. https://doi.org/10.1111/j.1525-1497.2005.40269.x.
23. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial intelligence (AI) applications for COVID-19 pandemic. Diabetes Metabolic Syndrome: Clin Res Reviews. 2020. https://doi.org/10.1016/j.dsx.2020.04.012.
24. Punitha S, Al-Turjman F, Stephan T. An automated breast cancer diagnosis using feature selection and parameter optimization in ANN. Comput Electr Eng. 2021. https://doi.org/10.1016/j.compeleceng.2020.106958.
25. Shah A, Shah M, Pandya A, Sushra R, Sushra R, Mehta M,… and Patel K. A comprehensive study on skin cancer detection using artificial neural network (ANN) and convolutional neural network (CNN). Clinical eHealth. 2023. https://doi.org/10.1016/j.ceh.2023.08.002.
26. Hung M, Li W, Hon ES, Su S, Su W, He Y,… and Lipsky MS. Prediction of 30-day hospital readmissions for all-cause dental conditions using machine learning. Risk Manage Healthcare Policy. 2020. https://doi.org/10.2147/RMHP.S272824.
27. Li W, Lipsky MS, Hon ES, Su W, Su S, He Y,… and Hung M. Predicting all-cause 90-day hospital readmission for dental patients using machine learning methods. BDJ Open. 2021; https://doi.org/10.1038/s41405-021-00057-6.
28. Tey SF, Liu CF, Chien TW, Hsu CW, Chan KC, Chen CJ,… and Wu WS. Predicting the 14-day hospital readmission of patients with pneumonia using artificial neural networks (ANN). Int J Environ Res Public Health. 2021. https://doi.org/10.3390/ijerph18105110.

29. Nas S, Koyuncu M. Emergency department capacity planning: a recurrent neural network and simulation approach. Computational and mathematical methods in medicine. 2019; https://doi.org/10.1155/2019/4359719.
30. Braga MDB, Fernandes RDS, Souza Jr GND, Rocha JECD, Dolácio CJF, Tavares Jr IDS,… and Vallinoto ACR. Artificial neural networks for short-term forecasting of cases,deaths, and hospital beds occupancy in the COVID-19 pandemic at the Brazilian Amazon. PloS one. 2021; https://doi.org/10.1371/journal.pone.0248161.
31. Triana AJ, Vyas R, Shah AS, Tiwari V. Predicting length of stay of coronary artery bypass grafting patients using machine learning. J Surg Res. 2021. https://doi.org/10.1016/j.jss.2021.02.003.
32. Kirchebner J, Günther MP, Sonnweber M, King A, Lau S. Factors and predictors of length of stay in offenders diagnosed with schizophrenia-a machine-learning-based approach. BMC Psychiatry. 2020. https://doi.org/10.1186/s12888-020-02612-1.
33. Iwase S, Nakada TA, Shimada T, Oami T, Shimazui T, Takahashi N,… and Kawakami E.Prediction algorithm for ICU mortality and length of stay using machine learning. Scie Rep. 2022. https://doi.org/10.1038/s41598-022-17091-5.
34. Morton E, Marzban G, Giannoulis A, Patel R, Aparasu R, Kakadiaris IA. A comparison of supervised machine learning techniques for predicting short-term ın-hospital length of stay among diabetic patients. :Detroit: 2014 13th International Conference on Machine Learning and Applications; 2014. pp. 428–431. https://doi.org/10.1109/ICMLA.2014.76.
35. Thompson B, Elish KO, Steele R, Machine Learning-Based Prediction of Prolonged Length of Stay in Newborns, 2018 17th IEEE International Conference on Machine Learning and, Applications. (ICMLA): Orlando; 2018. pp. 1454–1459. https://doi.org/10.1109/ICMLA.2018.00236.
36. Maharlou H, Kalhori SRN, Shahbazi S, Ravangard R. Predicting length of stay in intensive care units after cardiac surgery: comparison of artificial neural networks and adaptive neuro-fuzzy system. Healthc Inf Res. 2018;24(2):109–17.
37. Daghistani TA, Elshawi R, Sakr S, Ahmed AM, Al-Thwayee A, Al-Mallah MH. Predictors of in-hospital length of stay among cardiac patients: a machine learning approach. Int J Cardiol. 2019. https://doi.org/10.1016/j.ijcard.2019.01.046.
38. Castiñeira D, Schlosser KR, Geva A, Rahmani AR, Fiore G, Walsh BK,… and Santillana,M. Adding continuous vital sign information to static clinical data improves the prediction of length of stay after intubation: a data-driven machine learning approach. Respir Care. 2020; 65(9):1367–1377.
39. Kulkarni H, Thangam M, Amin AP. Artificial neural network-based prediction of prolonged length of stay and need for post‐acute care in acute coronary syndrome patients undergoing percutaneous coronary intervention. Eur J Clin Invest. 2021. https://doi.org/10.1111/eci.13406.
40. Wu J, Lin Y, Li P, Hu Y, Zhang L, Kong G. Predicting prolonged length of ICU stay through machine learning. Diagnostics. 2021. https://doi.org/10.3390/diagnostics11122242.
41. Fu Z, Gu X, Fu J, Moattari M, Zulkernine F. Predicting the length of stay of patients in hospitals. In 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 3150–3156). IEEE. 2021; https://doi.org/10.1109/BIBM52615.2021.9669527.
42. Abd-Elrazek MA, Eltahawi AA, Abd Elaziz MH, Abd-Elwhab MN. Predicting length of stay in hospitals intensive care unit using general admission features. Ain Shams Eng J. 2021;12(4):3691–702.
43. Orooji A, Shanbehzadeh M, Mirbagheri E, et al. Comparing artificial neural network training algorithms to predict length of stay in hospitalized patients with COVID-19. BMC Infect Dis. 2022;22:923. https://doi.org/10.1186/s12879-022-07921-2.
44. Pan Y, Fang C, Zhu X, Wan J. Construction of a predictive model based on MIV-SVR for prognosis and length of stay in patients with traumatic brain injury: retrospective cohort study. Digit Health. 2023. https://doi.org/10.1177/20552076231217814.
45. Namavarian A, Gabinet-Equihua A, Deng Y, Khalid S, Ziai H, Deutsch K, de Almeida JR. Length of stay prediction models for oral cancer surgery: machine learning, statistical and ACS-NSQIP. Laryngoscope. 2024;134(8):3664–72. https://doi.org/10.1002/lary.31443.
46. Erdogan Yildirim A, Canayaz M. Machine learning-based prediction of length of stay (LoS) in the neonatal intensive care unit using ensemble methods. Neural Comput Applic. 2024;36:14433–48. https://doi.org/10.1007/s00521-024-09831-7.
47. Naemi A, Schmidt T, Mansourvar M, et al. Quantifying the impact of addressing data challenges in prediction of length of stay. BMC Med Inf Decis Mak. 2021;21:298. https://doi.org/10.1186/s12911-021-01660-1.
48. Golafshani EM, Behnood A, Arashpour M. Predicting the compressive strength of normal and High-Performance concretes using ANN and ANFIS hybridized with grey Wolf optimizer. Constr Build Mater. 2020. https://doi.org/10.1016/j.conbuildmat.2019.117266.
49. Buber E, Şahingoz OK. Makine Öğrenmesi sistemi ile görüntü İşleme ve en uygun parametrelerin ayarlanması. In Artificial Intelligence and Data Processing Symposium (IDAP)2017. (pp. 16–17).
50. Yadav N, Yadav A, Kumar. M. An introduction to neural network methods for differential equations. New York: Springer; 2015.
51. Tolon M, Tosunoğlu NG. Analyzing consumer satisfaction data: a comparison of artificial neural networks and regression analysis Gazi Universitesi Iktisadi ve. Idari Bilimler Fakultesi Dergisi. 2008;10(2):247–259.
52. Ma L, Khorasani K. A new strategy for adaptively constructing multilayer feedforward neural networks. Neurocomputing. 2003;51:361–85.
53. Mouloodi S, Rahmanpanah H, Gohery S, Burvill C, Davies HM. Feedforward backpropagation artificial neural networks for predicting mechanical responses in complex nonlinear structures: A study on a long bone. J Mech Behav Biomed Mater. 2022;128:105079.
54. Kulkarni P, Watwe V, Chavan T, Kulkarni S. Artificial neural networking for remediation of methylene blue dye using fuller's earth clay. Curr Res Green Sustainable Chem. 2021;100131. https://doi.org/10.1016/j.crgsc.2021.100131.
55. Abiodun OI, Jantan A, Omolara AE, Dada KV, Mohamed NA, Arshad H. State-of-the-art in artificial neural network applications: A survey. Heliyon. 2018;4(11):e00938. https://doi.org/10.1016/j.heliyon.2018.e00938.
56. Esfe MH, Toghraie D, Amoozadkhalili F. Optimization and design of ANN with Levenberg-Marquardt algorithm to increase the accuracy in predicting the viscosity of SAE40 oil-based hybrid nano-lubricant. Powder Technol. 2023;415:118097.
57. Saini LM. Peak load forecasting using bayesian regularization, resilient and adaptive backpropagation learning based artificial neural networks. Electr Power Syst Res. 2008;78:1302–10.
58. Demuth H, Beale M. Neural network toolbox user's guide version 4; The Math Works Inc.: Natick, MA, USA, 2000; pp. 5–22.
59. Marquardt D. An algorithm for least-squares Estimation of non-linear parameters. J Soc Industrial Appl Math Soc Industrial Appl Math. 1963;11(2):431–41.
60. Khan N, Gaurav D, Kandl T. Performance evaluation of Levenberg-Marquardt technique in error reduction for diabetes condition classification. Procedia Comput Sci. 2013;18:2629–37.
61. Kayri M. Predictive abilities of bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: a comparative empirical study on social data. Math Comput Appl. 2016;21(2):20.
62. Burden F, Winkler D. Bayesian regularization of neural networks. Methods Mol Biol. 2008;458:25–44.
63. Okut H, Gianola D, Rosa GJM, Weigel KA. Prediction of body mass index in mice using dense molecular markers and a regularized neural network. Genet Res Camb. 2011;93:189–201.
64. Møller MF. A scaled conjugate gradient algorithm for fast supervised learning. Neural networks. Elsevier. 1993;6(4):525–33.
65. Apicella A, Donnarumma F, Isgrò F, Prevete R. A survey on modern trainable activation functions, Neural Networks, 2021;138:14–32, 2021. https://doi.org/10.1016/j.neunet.2021.01.026.
66. Laabid Z, Moumen A, Mansouri K, Siadat A. Numerical study of the Speed's response of the various intelligent models using the Tansig, Logsig and Purelin activation functions in different layers of artificial neural network. IAES Int J Artif Intell. 2023;12(1):155.
67. Hornik K, Stinchcombe M, White H. Multi-layer feed-forward networks are universal approximators. Neural Netw. 1989. https://doi.org/10.1016/0893-6080(89)90020-8.
68. Haykin S. Neural networks: a comprehensive foundation. Prentice Hall PTR; 1998.
69. Li X, Liu Z, Cui S, Luo C, Li C, Zhuang Z. Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning. Comput Methods Appl Mech Eng. 2019;347:735–53. https://doi.org/10.1016/j.cma.2019.01.005.
70. Karaboga D, Kaya E. Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. Artif Intell Rev. 2019. https://doi.org/10.1007/s10462-017-9610-2.
71. AlRassas AM, Al-Qaness MA, Ewees AA, Ren S, Abd Elaziz M, Damaševičius R, Krilavičius T. Optimized ANFIS model using Aquila optimizer for oil production forecasting. Processes. 2021. https://doi.org/10.3390/pr9071194.
72. Yılmaz A. Cancer risk analysis with using neuro-fuzzy logic model (Doctorate Thesis). Sakarya: Sakarya University; 2015.

73. Salleh MNM, Talpur N, Hussain K. Adaptive neuro-fuzzy inference system: overview, strengths, limitations, and solutions. In: Tan Y, Takagi H, Shi Y, editors Data mining and big data. DMBD. Lecture notes in computer Science. Cham: Springer; 2017, vol 10387. https://doi.org/10.1007/978-3-319-61845-6_52.
74. Hosseini MS. Review of medical image classification using the adaptive neuro-fuzzy inference system. J Med Signals Sens. 2012;2(1):49–60. ·Zekri M.
75. Yeom CU, Kwak KC. Performance comparison of ANFIS models by input space partitioning methods. Symmetry. 2018. https://doi.org/10.3390/sym10120700.
76. Uyanık GK, Güler N. A study on multiple linear regression analysis. Procedia-Social Behav Sci. 2013. https://doi.org/10.1016/j.sbspro.2013.12.027.
77. Su Y, Gao X, Li X, Tao D. Multivariate multilinear regression. IEEE Trans Syst Man Cybernetics Part B (Cybernetics). 2012. https://doi.org/10.1109/TSMCB.2012.2195171.
78. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. Volume 112. New York: springer; 2013. p. 18.
79. Gul M, Guneri AF. Forecasting patient length of stay in an emergency department by artificial neural networks. J Aeronaut Space Technol. 2015;8(2):43–8.
80. Dogu E, Albayrak YE, Tuncay E. Length of hospital stay prediction with an integrated approach of statistical-based fuzzy cognitive maps and artificial neural networks. Med Biol Eng Comput. 2021;59:483–96. https://doi.org/10.1007/s11517-021-02327-9.
81. Waller D. Operations management: a supply chain approach (No. hal-02297985). Cengage Learning Business Press London; 1999.
82. Cybenko G. Approximation by superpositions of a sigmoidal function. Math Control Signals Syst. 1989;2(4):303–14.
83. Hornik K, Stinchcombe M, White H. Multi-layer feed-forward networks are universal approximators. Neural Netw. 1989;2(5):359–66.
84. Sharma B, Venugopalan KJIJCE. Comparison of neural network training functions for hematoma classification in brain CT images. IOSR J Comput Eng. 2014;16(1):31–5.
85. Das A, Ben-Menachem T, Cooper GS, Chak A, Sivak MV, Gonet JA, Wong RC. Prediction of outcome in acute lower-gastrointestinal haemorrhage based on an artificial neural network: internal and external validation of a predictive model. Lancet. 2003. https://doi.org/10.1016/S0140-6736(03)14568-0.
86. LaFaro RJ, Pothula S, Kubal KP, Inchiosa ME, Pothula VM, Yuan SC,… and Inchiosa Jr MA. Neural network prediction of ICU length of stay following cardiac surgery based on pre-incision variables. PLoS One. 2015; https://doi.org/10.1371/journal.pone.0145395.
87. Gul M, Guneri AF. Planning the future of emergency departments: forecasting ED patient arrivals by using regression and neural network models. Int J Industrial Eng. 2016;23(2):137–54.
88. Jamei M, Nisnevich A, Wetchler E, Sudat S, Liu E. Predicting all-cause risk of 30-day hospital readmission using artificial neural networks. PLoS ONE. 2017. https://doi.org/10.1371/journal.pone.0181173.
89. Klemt C, Tirumala V, Barghi A, Cohen-Levy WB, Robinson MG, Kwon YM. Artificial intelligence algorithms accurately predict prolonged length of stay following revision total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc. 2022. https://doi.org/10.1007/s00167-022-06894-8.
90. Turkish Republic Personal. Data Protection Law. https://www.mevzuat.gov.tr/mevzuatMevzuatNo=6698&MevzuatTur=1&MevzuatTertip=5.

Publisher's note