Predictive Model of Life Cycle Medical Care Services in Peruvian Health Private Entities

Juan Carlos Herrera Miranda[1], Javier Mendoza-Montoya[2], Jorge Victor Mayhuasca Guerra[1] and Juan Jesús Soria Quijaite[2]

[1] Universidad Nacional Federico Villareal, Jr. Carlos Gonzáles 285 Urb. Maranga-San Miguel, Lima, Perú
[2] Universidad Tecnológica del Perú, Av. Parra 203, Arequipa, Perú

Abstract

A predictive model is a good technique for prognosticating the life cycle of medical care services (LC-MCS) related to its growth, stability, and decline to support information processes and decision-making based on intelligent forecasts. Medical care services require organization, scheduling, and programming attention to prevent the capacity of medical personnel. Otherwise, it could be chaotic. This research aims to evaluate three predictive models to find the best one to fit information about the LC-MCS and understand medical care services' behavior in healthcare entities. The proposed model applied to Ricardo Palma Clinic (RPC) in Lima Perú analyzes the LC-MCS information based on 4950 clinical health records. Three predictive models were trained and compared to evaluate the accuracy of backpropagation neural networks, decision trees, and multiple linear regression models. After training, the coarse tree model gives a root mean squared error (RMSE) of 30.237 and an accuracy of 85%. The neural network model with ten hidden layers using the Sigmoid transfer function gives the best validation performance of 608083 at epoch 9; however, the Stepwise Linear regression model gives the best performance between the three trained models with a RMSE of 11.553 and 87.3% accuracy in predicts LC-MCS in Ricardo Palma Clinic. In conclusion, it is possible to predict the LC-MCS in Peruvian Healthcare entities and use tools such as stepwise linear regression to give real-time information about medical care services.

Keywords

Healthcare predictive model, linear regression, life cycle medical service, neural network model, decision tree1

1. Introduction

Due to the state's limited capacity to deliver high-quality healthcare to its population, the segmented structure and organization of the Peruvian health system show us the inequity of healthcare services in Perú, which has a population of about 33.72 million people, according to the INEI report [1]. Public and private health entities support the Peruvian healthcare system and generate different funding sources, variate insurance schemes, and multiple health service attention channels. The Ministry of Health (MINSA) oversees the beneficiary population based on the state's general taxes that supports the integral health insurance program (SIS) that covers insurance for poor and vulnerable people (63.5%). Ministry of Defense finances health insurance for personnel of army forces (0.5%) and police force (1.5%); EsSalud provides health care, pension, and welfare coverage for (29.3%) of the population; payroll workers finance that and is dependent on the labor Ministry[2].

Many private healthcare entities provide insurance and cover a small fraction of the population (9.6%), with some overlap with EsSalud; however, Peru does not currently have a system that integrates and links health information between its public and private entities, but many applications exist that

CITIE 2023: International Congress on Education and Technology in Sciences, December 04–06, 2023, Zacatecas, Mexico
jcherreramiranda@mail.com (J.C. Herrera Miranda); jmendozamo@utp.edu.pe (J. Mendoza-Montoya); jmayhuasca@unfv.edu.pe (J. V. Mayhuasca Guerra); c20723@utp.edu.pe (J. J. Soria Quijaite)
0000-0002-5640-400X (J.C. Herrera Miranda); 0000-0001-9365-1723 (J. Mendoza-Montoya); 0000-0002-6465-4738 (J. V. Mayhuasca Guerra); 0000-0002-4415-8622 (J. J. Soria Quijaite)

provide segmented information in public and private health systems; for this reason, interoperability among these systems is minimum or inexistent because of their standards and protocols [2]. Recently electronic medical reports (EMRs) are getting more interest in private entities because of the easy way to manage the information and improve hospital care health services [3], [4].

Prediction models are innovative machine learning techniques that support processing and analytical data, especially in health care services based on EMRs or electronic health records (EHRs); the continued growing amount of data is hard to manage and analyze without the support of artificial intelligence and machine learning models [5]–[7]. The application of predictive models with machine learning algorithms in medical services is increasing worldwide because of its great importance as a support to improve healthcare attention; in some cases, predictive models can predict diseases in a short time diagnostic and with reduced costs [8]–[13]. Recently, many prediction models applied to the health sector were developed with machine learning models to identify chronic diseases in their earlier stages bringing support to doctors in determining sick conditions accurately and therefore minimizing the patient risk factor that leads to death. In health care some prediction models have been used with good results, to mention just a few cases: early detection of late-onset neonatal sepsis with machine learning approach, mild cognitive impairment using random forest model, infection disease spread based on migration and effective distance using linear regression model, and decision tree model to predict the outcome of non-intensive hospitalized for COVID-19 [14]–[20].

Therefore, healthcare information management with machine learning models is essential to doctors because of supporting effective diagnosis to determine in patients their bad o favorable disease condition evolution[21]–[23]. It is important to value demographic data, medical analysis reports, and patient's clinical history because they are essential to feeding machine learning models and training with appropriate algorithms could bring solutions in early-stage disease diagnosis and improve inpatient attention based on predictive information of the LC-MCS and provide sufficient medical personnel in healthcare entities.

Research in prognosis diseases with predictive models is extensive, and there are many applications related that supports medical staff to give more precise diagnostics; although, into medical care service life cycle predictions there is not enough applications, however they are essential for evaluating the medical care services in public and private entities; therefore, in this work, we propose a study of three predictive models to determine the life-cycle of medical care services in Peruvian private health entities, taking the case of Ricardo Palma Private Clinic that has 14 health service branches in different districts of Lima-Perú. The main objective of this research is to predict the growth, stability and decline of medical care services to evaluate their processes and provide them with a predictive tool that allows managers to take decisions that can optimize their medical care services and provide a fast response to the population. Thus, in this research we propose and evaluate the response of three predictive models: decision tree model, neural network, and linear regression models; after their training a comparison between the three model will be made to determine the models with the best performance.

The organization of this article is as follow: Section II presents the related works addressed to health diseases prediction. Section III describes the methodology. Section IV presents the results of the three models in details and data analysis proposes. Section V presents the discussion of the proposed three models, and Section VI presents the conclusions.

1.1. Related works

During the COVID-19 pandemic disease, people were afraid to move out their houses and traveling was restricted around all the world; in this scenario, many researchers tried to find different predictive models to control and stop the pandemic propagation into the population. An interesting contribution proposed by T.Zhou [16] show us the application of linear regression model to analyze the relationship between the effective distance between a city labeled as a focus of contagion and its neighboring cities, the epidemic transmission trajectory, and the arrival time of the COVID-19 in Wuhan, the model has a good fit with the level of cumulative of confirmed cases; therefore, it can be helpful as a reference of early infection diseases prevention in future pandemic events.

A machine learning approach with a decision tree model was used to analyze the clinical data of hospitalized COVID-19 patients to determine an efficient predictive model based on clinical, demographic, and blood chemistry parameters of the patients to predict three possible outcomes: discharged alive, transferred to ICU or death, whichever occurs first. The decision tree model obtained 75.93% of accuracy and a sensitivity of 99.61% demonstrating that this model can support medical staff in take prevision with inpatients according to its health evolution [17].

Electronic Health Record (EHR) have rapidly advanced last years for predictive analytics into clinical settings for disease evolution stratification. Most studies were conducted in inpatient academic settings. Some cases include alert fatigue, lack of exercise, and increased work burden on the care team. Of 32 studies that reported effects on clinical outcomes, 22 (69%) demonstrated improvement after model implementation. Overall, EHR-based predictive models offer promising results for improving clinical outcomes, although several gaps in the literature remain, and most study designs were observational. Future studies can be applied to medical care services predictions to enhance clinical management [5].

2. Methodology

This study is focused as a quantitative research and descriptive achievement regarding the three predictive models trained to determine the life cycle of medical care services in Ricardo Palma Clinic, with the aim to improve the actual information of inpatients medical services; therefore, the method was divided into three stages: data source, data processing, and data evaluation.

2.1. Data Source

The sources of data are supported by inpatients clinical health record (CHR) provided by the manager of the RPC clinic with the aim to analyze the actual information, the clinical records was selected regarding about 12578 CHR of inpatients treated on 2017 period in the RPC taking into account the main clinic and the 14 health service branches distributed in most populated districts of Lima-Perú. The sample size was calculated with equation (1) based on limit theorem of conventional statistic method and performed with the parameters described in Table 1. The sample size (n) calculated is 372, but this theoretical value has been adjusted in the range of 299 to 375 samples between the 14 health establishments belonging to the 14 districts most populated; therefore, the total number of samples add up to 4590 clinical health record of inpatients.

$$n = \frac{N.Z^2.p.q}{e^2(N-1)+Z^2.p.q} \quad (1)$$

Table 1
Sample of Clinical Health Records

Description	Variable	Value
Population size	**N**	**12 578**
Significancy level	**α**	**0.05**
Confidence level	**1-α**	**0.95**
Z -Score	**Z(1-α)**	**1.96**
Favorables cases	**p**	**0.5**
Desfavorables cases	**q=1-p**	**0.5**
Precision error	**e**	**0.05**

Minimun Sample size	**n**	**372**
Sample size		**4590**

The total number of CHR of inpatients used in this study goes to 4590, they are distributed among 22 medical care services treated at the RPC. In figure 1, is shown all medical services presented in descending order listed from the most required health services to the even the least required for the population; also, we can denote that Allergies health service are the most required by population in Lima-Perú, maybe because of the high intensity of UV radiation and for the air pollution that people is exposed every day; then, follows mastology, oncologic surgery, neurosurgery, cardiology, nephrology, gynecology psychiatrist, dermatology, neurology, chest surgery, clinical lab, gastroenterology, endocrinology, and pediatrics report over 200 inpatients. The last seven health services: urology, ultrasound, ophthalmology, geriatric, radiology, traumatology, and infectology are also required, but just in less quantity.

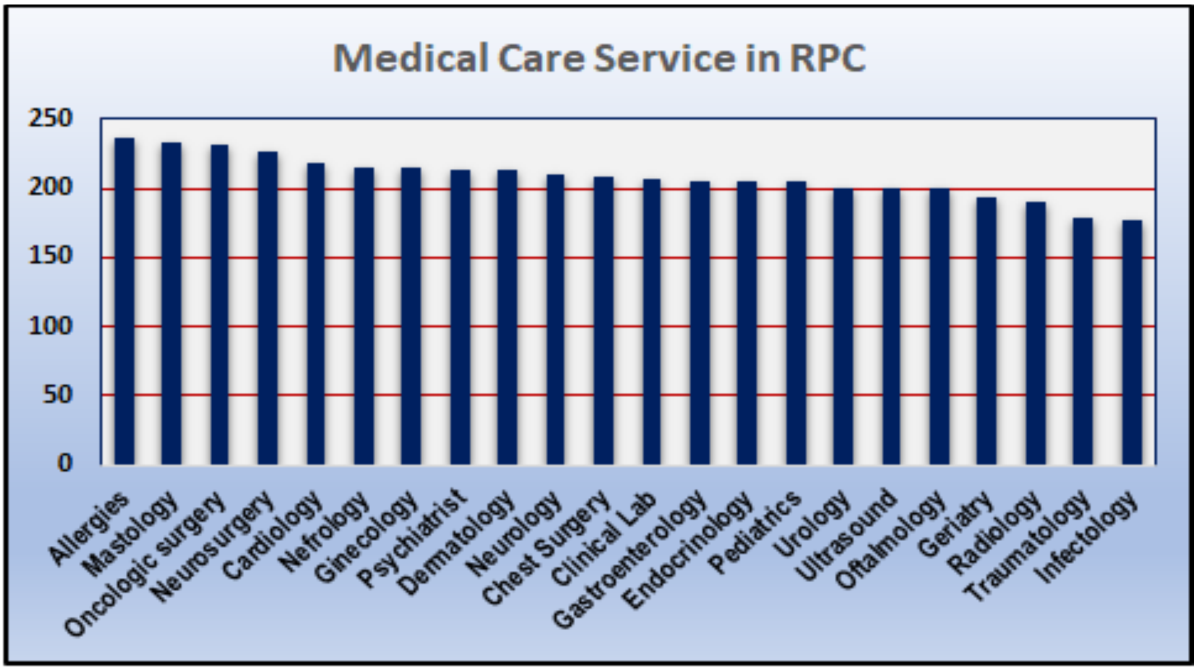

Figure 1: List of medical care services in RPC

In Table 2, is shown the total sample within the 4950 inpatients CHR and distributed in five different age groups to get the most related and precise information of life cycle medical care service.

Table 2
Outpatients by age group

District	Children	Teenage	Young	Adult	Elderly	Subtotal
Chorrillos	**85**	**77**	**63**	**78**	**72**	**375**
Lince	60	79	62	70	80	351
San Borja	71	73	66	66	73	349
Surquillo	67	73	68	69	67	344
La Molina	59	59	71	81	67	337
Lima	72	58	62	63	73	328
Pueblo Libre	54	65	76	59	70	324

Comas	69	72	50	62	70	323
Breña	69	60	67	66	58	320
Surco	59	68	54	65	72	318
San Luis	54	67	60	68	65	314
Jesús María	67	59	57	64	58	305
San Isidro	73	52	62	52	64	303
Miraflores	69	61	56	51	62	299
Total	928	923	874	914	951	4590

Regarding the Table 2, the total number of CHR by age group is as follow: 928 inpatients belong to children age group, 923 to teenage, 874 to young, 914 to adult, and 951 to elder inpatients; also, the mean value of the number of CHR by district is 354, however the sample size of every district oscillates between 299 and 375 CHR; finally the 4950 CHR samples are distributed in 14 districts, and every district in five age group, and every age group in 22 medical care service.

The importance of analyze the data separately by every district can give us particular information of the outpatients, for example in Chorrillos district, Lince, and San Borja have a greater number of elderly CHR; in contrast to Jesus María, San Isidro, and Miraflores that have minor number of elderly CHR, this difference is an important point to find relationship between medical health services and any district. Other criteria to evaluate data by district is to determine the evolution of increase or decrease inpatients in all medical care services and evaluate which services are growing or declining in the course of time.

2.2. Data Processing

To process the data collected of the 14 districts of Lima, three predictive models were applied and trained for processing the 4590 CHR: a Backpropagation neural network (ANN), a decision tree, and a multivariate regression model were trained and tested to predict the life cycle of the 22 medical care services in RPC. The interface was implemented in Power BI, an open access platform that facilitates the graphical presentation of the historical behavior of inpatients in RPC.

The block diagram in Fig. 2 shows us the three basic steps to follow in getting the predictive model of life cycle medical care service: at the first block we collect a dataset of 4950 CHR that belongs to 22 medical care services, then we train the three predictive models to obtain the best accuracy, and finally results are presented in graphics and evaluated with different metrics.

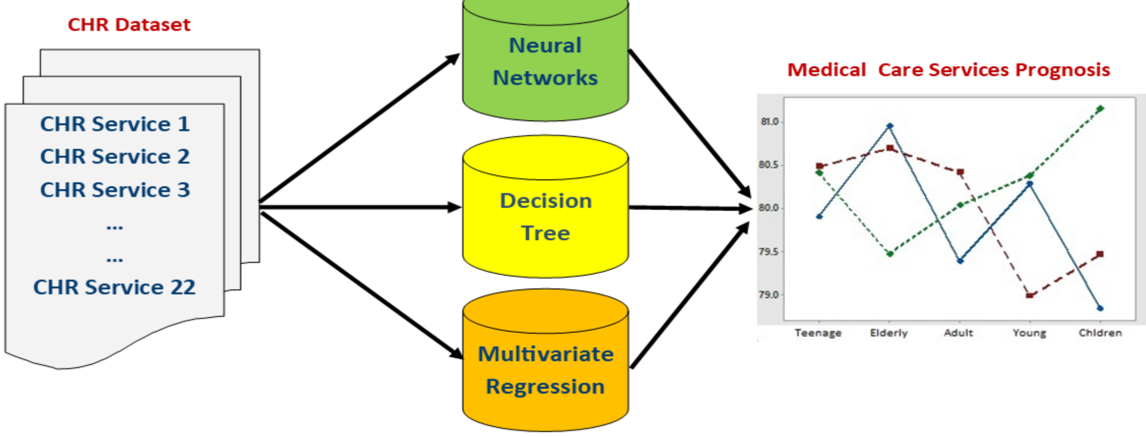

Figure 2: Predictive models applied to healthcare services

2.2.1. Backpropagation Neural Network

Neural networks process information and is inspired in the human nervous system; thus, artificial neuron networks are mathematical functions derived from biological neurons. The block diagram of the artificial neural network model is shown in Figure 3, where Xi = X1,X2,X3…Xn, and represents the input data; Wi = W1,W2,W3…Wn, represents the weight assigned to a every input data; to offset the results usually is used a Bias which is added to the product of features and weights; then, the summing junction biased activate the desired function and gives out the output response.

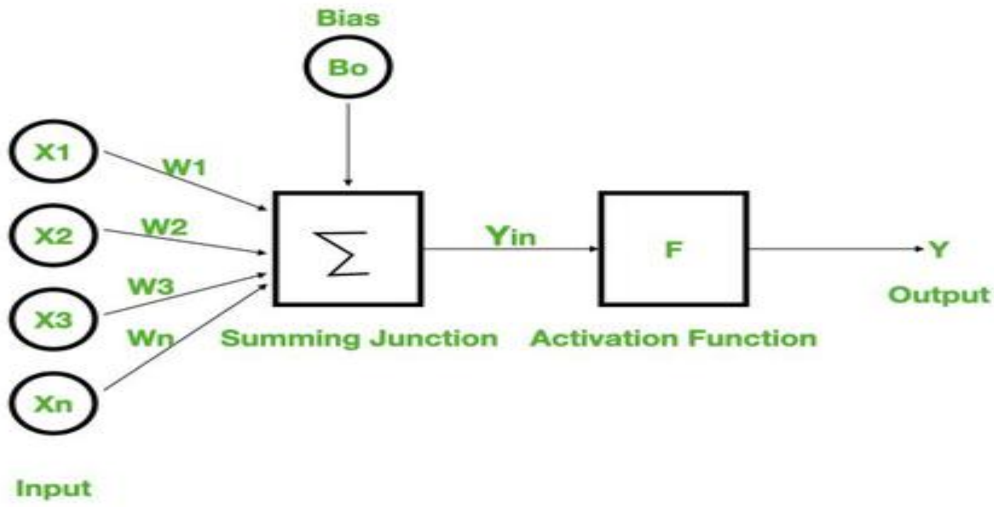

Figure 3: Artificial neural network representation

The back propagation algorithm applied to a neuronal network is described below:

1. Input Xi, arrive to the preconnected layer.
2. The input is modeled with a defined weight according to the requirements.
3. Calculate the output of each neuron from the input layer to the hidden layer and to the output layer.
4. Calculate the error in the outputs, where backpropagation error = Actual error – desired output.
5. From the output layer, go back to the hidden layer to adjust the weights to reduce the error.
6. Repeat the process until the desired output is achieved.

The metric used to measure the performance on backpropagation nets is usually the mean square error (MSE) defines with the equation (2):

$$E = \frac{1}{n} . \sum_{i=0}^{n} [\sum_{j=0}^{m} (Y_{ij} - O_{ij})^2] \quad (2)$$

Where Yij represent the correct outputs for a pattern I; Oij represent the network estimates for pattern i; m is the number of output nodes; and n is the number of training patterns. These criteria is similar to least square error used in classical regression analysis. Therefore, backpropagation learning can be a generalization of classical regression, a sort of super regression.

2.3. Decision tree

This algorithm works as a flow chart structure, where the root node is the top node in a decision tree; the internal nodes represent any attribute, the branch represent a decision rule and each leaf node represent the outcome. Our model will use this algorithm to predict the medical care service in RPC processing the 4950 CHR classifying 22 medical care services for the 14 RCP clinic branches in Lima Perú.

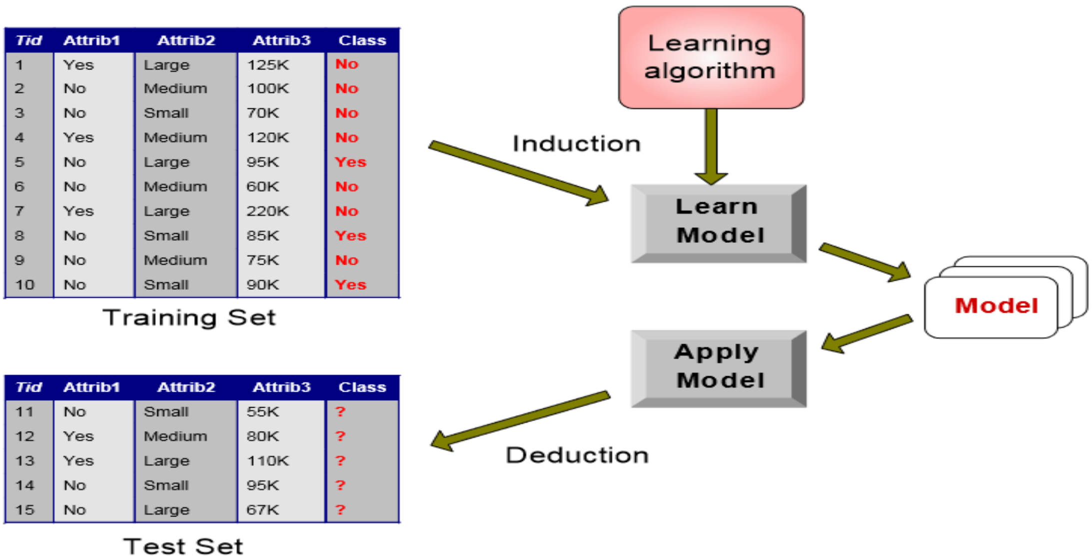

Figure 4: Tree decision model representation

2.4. Multivariate Regression

Is this model the dependent variable (Y) represent the predicted value as a response of the multivariate equation that is composed of the summatory of (Xn) independent variables with its respective constant coefficient (β_n) that gives the proportional weight as a partial contribution of the total value prediction in the multivariate regression model; the aleatory error is represented by the (μ) variable and is essential to calculate the accuracy of the model. The equation (3) shows us the relationship between Y and Xn variables; also βn and μ error

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 \ldots \ldots \beta_n X_n + \mu \qquad (3)$$

Equation (2) can be expressed in matrixial mode as follow:

$$Y = \beta.X + \mu$$

An important metric to evaluate if any correlation exits in the relationship of Y and Xn variables is the F-Test that measure the mean square due to regression (MSR) over the mean square error (MSE); therefore, The F ratio represents a distribution with one numerator degree of freedom and n-2 denominator degrees of freedom. For this reason, it is often referred to as the analysis of variance F-test, taking as input data the sum of square residual (SSR) and the sum of square error (SSE). In Table 3 is shown the standard analysis of F-Test variance.

Table 3
F-Significancy test

Source	Sum of squares	Freedom degree	Mean Square	F-Test
Regression	**SSR**	**p**	**MSR = SSR/p**	**F=MSR/MSE**
Error	SSE	n-p-1	MSE = SSE/(n-p-1)	F=MSR/MSE
Total	SST	n-1		

The three model mentioned before will be under a trained process and the comparative results could let us choice the best model to predict the medical health care service of the 22 most required health care services for population in Lima-Perú along by its 14 RCP clinic branches; therefore, the goal of this is research is to demonstrate the following hypothesis: “Implementing a predictive model could improve the information about the life cycle of medical care services prognosis in Peruvian health entities". To probe and validate the formulated hypothesis, the following steps must be performed:

1) Make an initial diagnostic of the current situation to find actual indicators of the medical care services life cycle in RPPC without predictive models' application.
2) Apply three predictive models to compare each other between neuronal network, decision tree, and multivariate regression.
3) Determine the best predictive model that give us reliability in the results about medical care services life cycle in RPPC.

3. Results

The clinical health records sample belong to the dataset of 4590 patients attended in Ricardo Palma Clinic including its 14 clinic branches covering the principal districts of Lima. The medical health service records are performed in three quarters per year. The first quarter (Q1) registers 776 men and 707 women, totaling 1483 inpatients; the second quarter (Q2) registers 766 men and 750 women, totaling 1516 inpatients; and the third quarter (Q3) registers 820 men and 771 women, totaling 1591 inpatients; thus, 22 medical health services were required, where 2362 medical services were registered for men and 2228 for women. Clinical health records including 22 medical services have been tabulated by quarters and sorted in ascending order according to the most required services as we can see in Figure 5.

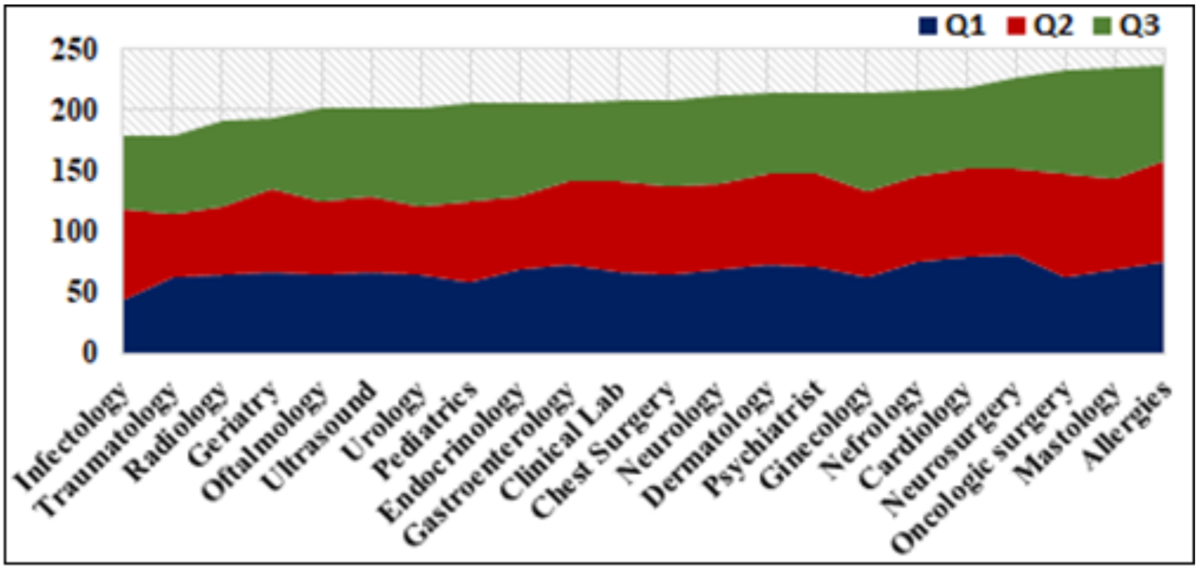

Figure 5: Quarterly inpatients on specialty medical services

Another point to analyze in Fig. 5 is the slight increase in the number of medical attentions as follows: first quarter 1483, second quarter 1516, and in third quarter 1591 respectively. However, is important to remark that the most requested medical services belong to allergies, mastology, oncology surgery, neurosurgery, and cardiology; therefore, in the opposite side infectology, traumatology, and radiology are the least requested medical services in RPC.

Now, the result of the predictive model is presented with comparative evaluations of different models that give us the chance to choose the first one with the best performance in decision tree, neural networks, and linear regression.

3.1. Predictive Model with Decision Tree

The first predictive model evaluated in our work with decision tree was performed with fine tree, medium tree, and coarse tree predictive models. The comparative results about decision tree models are shown in Table 4, where the comparative results of three predictive decision tree models are presented with four metrics to compare: root mean square error (RMSE), mean square error (MSE), mean absolute error (MAE), and the determination coefficient R2. The first model evaluated was Fine-Tree predictive model, and its metrics give us a root mean an RMSE of 36.158, an MSE of 1307.4, an MAE of 29.84 and a determination coefficient R2 of 0.57; it means that this statistical model predicts the outcome with 57% of precision, therefore it is considered as low performance approximation model; in second place, the Medium-Tree predictive model has an RMSE of 32.523, an MSE of 1057.7, an MAE of 27.32, and a determination coefficient R2 0f 0.62, which indicates a low performance in predicting an outcome with 62% precision, so on is not recommendable for our purpose; finally, the Coarse-Tree predictive model with RMSE of 30.237, a MSE of 914.3, a MAE of 25.84 and a determination coefficient R2 of 0.85, represents the highest performance between the three models, and its precision of 85% in prediction an outcome give us greater confidence in predict an outcome; therefore we choose Coarse-Tree predictive model as the best one with high probability to predict information about the life cycle of medical care services in private Peruvian health entities, specifically the Ricardo Palma Clinic of Lima. Table 4 resume the comparative results.

Table 4
Comparative results in decision tree models

Model	RMSE	MSE	MAE	R2
Fine-Tree	**36.158**	**1307.4**	**29.84**	**0.57**
Medium- Tree	32.523	1057.7	27.32	0.62
Coarse-Tree	**30.237**	914.3	25.84	0.85

3.2. Predictive Model with Neural Network

The implemented neuronal network predictive model is composed of 10 hidden layers and a sigmoidal function activation. Since 4950data were input in the neuronal network model to find the validation performance, 70% of the records were used in training the model, 15% for validation, and 15% for test the neuronal network performance. In Fig. 6 we can see how the best validation performance reaches 6080.8298 at epoch 6, where the training, validation and dispersion error had the lowest value that compared with the best tendency.

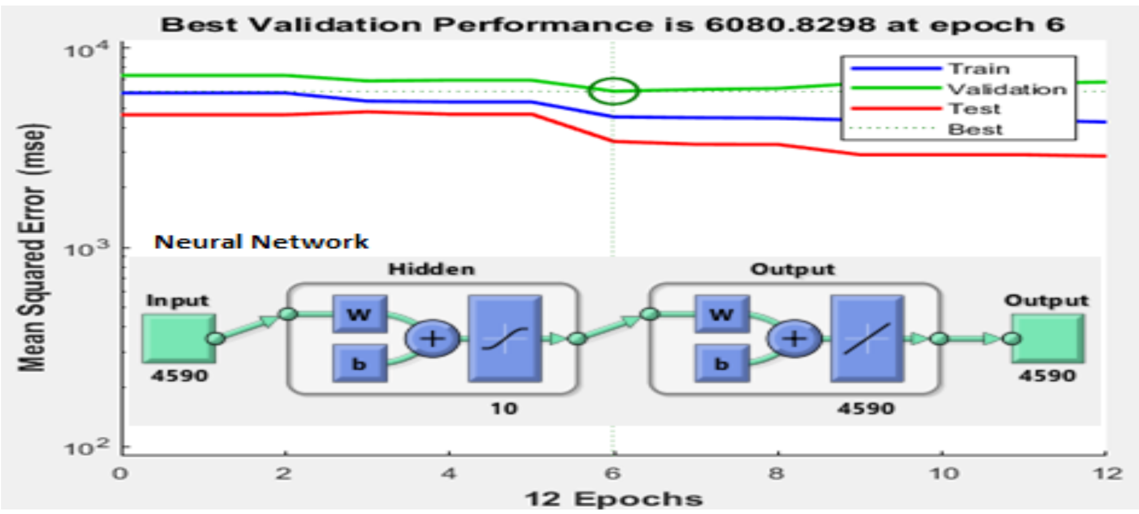

Figure 6: Neural network predictive model validation

Also is important to calculate the gradient, because this parameter measures the change in all weights respect to the changes in error, the gradient in a neuronal network represents the generalization of the derivative as a local slope of the multivariate functions and allow us to predict the effect of taking a small step from a point in any direction. In Fig. 7 the gradient and validation of the predictive model is shown, the gradient value obtained in epoch 12 is 17275.3308, with this value the neuronal network model is trustworthy to predict life cycle medical care services; therefore, the results give with a great confidence.

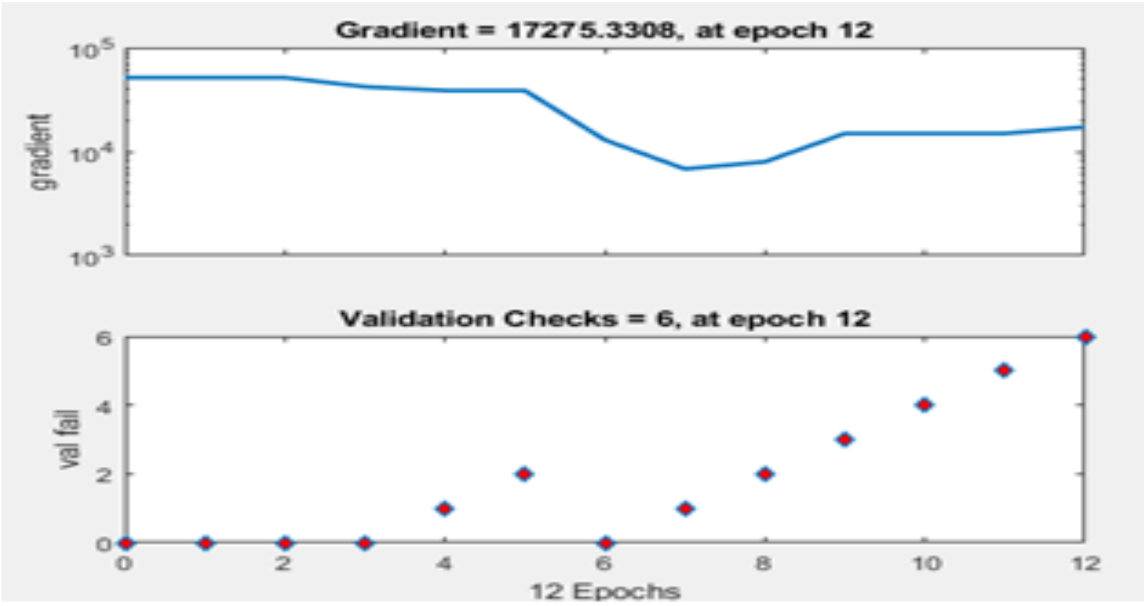

Figure 7: Measuring the gradient and validation in predictive model.

1.1 Predictive Model with Linear Regression

The linear regression predictive model belongs to supervised machine learning, its purpose is predicting the outcome based on the independent variable data points. The proposed linear regression model calculates the output with a minus slope of 0.0018 and a deviation of 48; the trained model is plotted and shown in Figure 8.

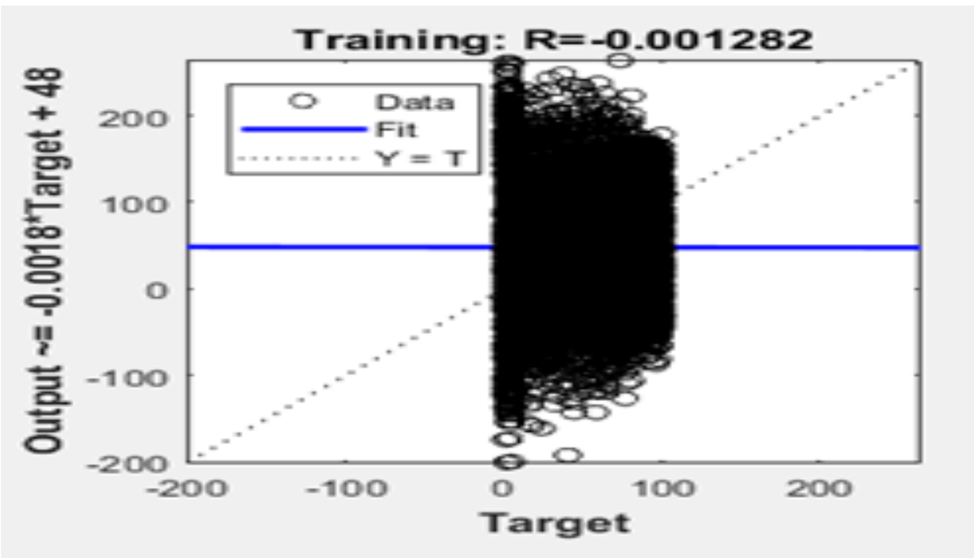

Figure 8: Linear Regression Training Model

Trained model plotted in Figure 8, can also be expressed with the following relationship

$$Output = -0.0018 * Target + 48 \quad (4)$$

The predictive model is processed with 4950 clinical health records of RPC inpatients belonging to 22 medical services and 14 districts of Lima, the training and validation takes a long time to process and calculate the predictor indicator of the life cycle medical service. Many proofs were made before obtaining the prediction metrics of the four linear regression models proposed to evaluate its accuracy and precision. Robust linear regression is less sensitive to outliers to standard linear regression, while interaction linear assess the relationship between each independent variable and the dependent variable, the stepwise linear regression is a model that iteratively examines the statistical significance of each independent variable.

Table 5
Comparative results between linear regression models

Model	RMSE	MSE	MAE	R^2
Standard Linear	**11.559**	**133.6**	**10.017**	**0.72**
Interaction Linear	11.746	137.97	9.92	0.714
Robust Linear	11.564	133.73	9.74	0.736
Stepwise Linear	11.553	133.47	9.43	0.873

The comparative metrics of linear regression models shown in Table 5 give us the tool to analyze the accuracy of the four evaluated models with metrics based on root mean square error (RMSE), main square error (MSE), main absolute error (MAE), and the coefficient of determination (R2) that will measure the performance of each linear regression model concerning to life cycle medical care services. The predictive model with the best performance is Stepwise linear regression, because its RMSE with 11.553 is the lowest between the four models and a coefficient of determination of 0.873 equivalent to 87.3% of precision in predicts the lifecycle of medical services; then continue in descendent order, robust linear regression with 73.6%, standard linear regression with 72%, and interaction linear regression with 71.4%. Finally, the total records are plotted and contrasted with the linear regression model with confidence level of 5% and significance level of 95%

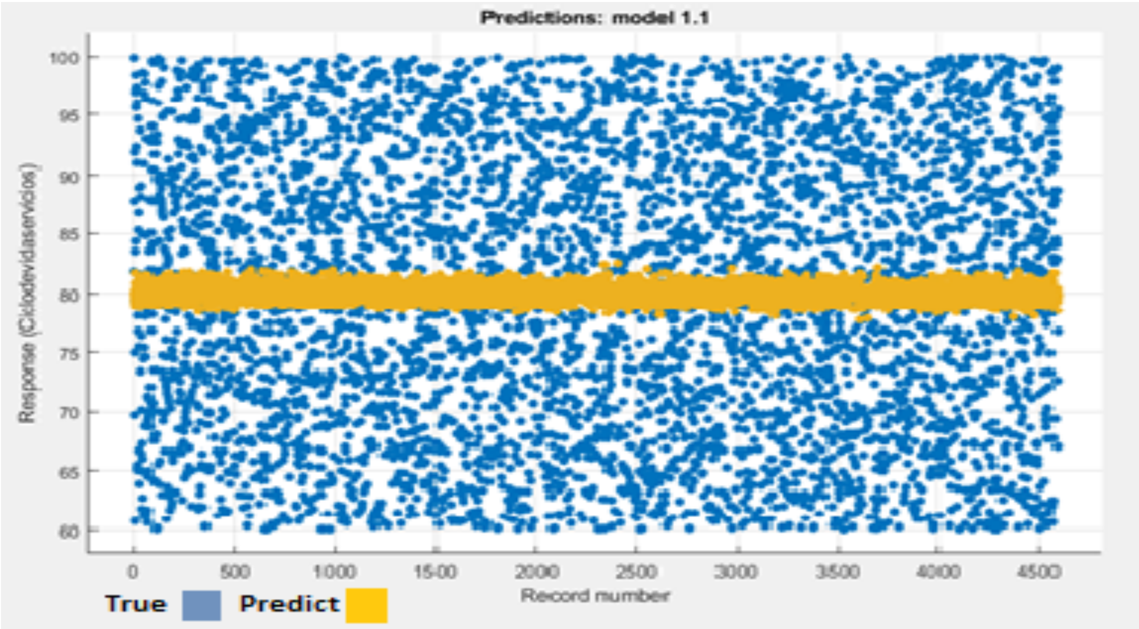

Figure 9: Predicted linear regression model.

Finally, the prediction model is represented in Figure 9, where the predicted linear regression model shows us the predict average response over the true data processed.

4. Conclusions

In this paper we have trained three predictive models: decision tree, neuronal network, and linear regression using 4950 clinical health records with the purpose of predict the LC-MCS and concluding that these predictive models improve the information of the LC-MCS in private entities of Peruvian health sector.

The analysis of MCS in Ricardo Palma clinic with a decision coarse-tree model give us the best performance with 85% of accuracy and 30.237 of RMSE; the neuronal network give us a best validation of 6068.8298 at epoch 6; however, the stepwise linear regression model give us the best global performance between the three models with a RMSE of 11.553 and 87.3% of accuracy in data prediction of LC-MCS.

As a future application with the stepwise linear regression model, we recommend analyze the 100% of clinical health records of RCP to obtain essential information to improve the management of medical care services and give this tool for well decision-making in Peruvian health private entities.

References

[1] INEI, "Perú: Estado de la Población en el año del bicentenario, 2021.," Inst. Nac. Estadística e informática, p. 76, 2021, [Online]. Available: https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1803/libro.pdf

[2] R. M. Carrillo-Larco et al., "Peru – Progress in health and sciences in 200 years of independence," Lancet Reg. Heal. - Am., vol. 7, no. December 2021, 2022, doi: 10.1016/j.lana.2021.100148.

[3] M. D. Sung et al., "Event prediction model considering time and input error using electronic medical records in the intensive care unit: Retrospective study," JMIR Med. Informatics, vol. 9, no. 11, pp. 1–12, 2021, doi: 10.2196/26426.

[4] A. Manuscript, "NIH Public Access," vol. 26, no. 3, pp. 246–251, 2010, doi: 10.1111/j.1471-1842.2009.00857.x.Biomedical.

[5] A. R. Kwilas, R. N. Donahue, K. Y. Tsang, and J. W. Hodge, “乳鼠心肌提取 HHS Public Access,” Cancer Cell, vol. 2, no. 1, pp. 1–17, 2015, doi: 10.3390/informatics7030025.Clinical.
[6] D. E. Leisman et al., “Development and Reporting of Prediction Models: Guidance for Authors From Editors of Respiratory, Sleep, and Critical Care Journals,” Crit. Care Med., vol. 48, no. 5, pp. 623–633, 2020, doi: 10.1097/CCM.0000000000004246.
[7] E. W. Steyerberg et al., “Assessing the performance of prediction models: A framework for traditional and novel measures,” Epidemiology, vol. 21, no. 1, pp. 128–138, 2010, doi: 10.1097/EDE.0b013e3181c30fb2.
[8] C. Krittanawong et al., “Machine learning prediction in cardiovascular diseases: a meta-analysis,” Sci. Rep., vol. 10, no. 1, pp. 1–11, 2020, doi: 10.1038/s41598-020-72685-1.
[9] C. Rogerson and M. Hall, “Methodological progress note: Machine learning methods in healthcare research,” J. Hosp. Med., no. March, pp. 431–434, 2023, doi: 10.1002/jhm.13078.
[10] A. Burlea-Schiopoiu and K. Ferhati, “The managerial implications of the key performance indicators in healthcare sector: A cluster analysis,” Healthc., vol. 9, no. 1, pp. 1–20, 2021, doi: 10.3390/healthcare9010019.
[11] A. Ramesh, C. K. Subbaraya, and G. K. Ravi Kumar, “AI based Dynamic Prediction Model for Mobile Health Application System,” Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 1, pp. 356–365, 2023, doi: 10.14569/IJACSA.2023.0140138.
[12] R. Alanazi, “Identification and Prediction of Chronic Diseases Using Machine Learning Approach,” J. Healthc. Eng., vol. 2022, 2022, doi: 10.1155/2022/2826127.
[13] H. Kwon, H. H. Kim, J. An, J. H. Lee, and Y. R. Park, “Lifelog data-based prediction model of digital health care app customer churn: Retrospective observational study,” J. Med. Internet Res., vol. 23, no. 1, pp. 1–12, 2021, doi: 10.2196/22184.
[14] H. F. Rammal and A. Z. Emam, “Heart failure prediction models using Big Data techniques,” Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 5, pp. 363–371, 2018, doi: 10.14569/IJACSA.2018.090547.
[15] E. B. Ing, “Neural network and logistic regression predictive calculator for giant cell arteritis,” Arch. Soc. Esp. Oftalmol., vol. 94, no. 12, p. 622, 2019, doi: 10.1016/j.oftal.2019.07.011.
[16] T. Zhou, “A Linear Regression Prediction Model of Infectious Disease Spread Based on Baidu Migration and Effective Distance,” Comput. Math. Methods Med., vol. 2022, no. December 2019, 2022, doi: 10.1155/2022/9554057.
[17] M. Giotta et al., “Application of a Decision Tree Model to Predict the Outcome of Non-Intensive Inpatients Hospitalized for COVID-19,” Int. J. Environ. Res. Public Health, vol. 19, no. 20, 2022, doi: 10.3390/ijerph192013016.
[18] H. Byeon, “A Prediction Model for Mild Cognitive Impairment Using Random Forests,” Int. J. Adv. Comput. Sci. Appl., vol. 6, no. 12, pp. 8–12, 2015, doi: 10.14569/ijacsa.2015.061202.
[19] Q. Wu et al., “Deep Learning Methods for Predicting Disease Status Using Genomic Data HHS Public Access,” J Biom Biostat, vol. 9, no. 5, 2018, [Online]. Available: http://creativecommons.org/licenses/by/4.0/:.
[20] W. Song, S. Y. Jung, H. Baek, C. W. Choi, Y. H. Jung, and S. Yoo, “A predictive model based on machine learning for the early detection of late-onset neonatal sepsis: Development and observational study,” JMIR Med. Informatics, vol. 8, no. 7, pp. 1–16, 2020, doi: 10.2196/15965.
[21] R. D. Riley and G. S. Collins, “Stability of clinical prediction models developed using statistical or machine learning methods,” pp. 1–30, 2022, [Online]. Available: https://arxiv.org/abs/2211.01061v1
[22] L. Belbasis and O. A. Panagiotou, “Reproducibility of prediction models in health services research,” BMC Res. Notes, vol. 15, no. 1, pp. 1–5, 2022, doi: 10.1186/s13104-022-06082-4.
[23] T. Davenport and R. Kalakota, “The potential for artificial intelligence in healthcare,” Futur. Healthc. J., vol. 6, no. 2, pp. 94–98, 2019, doi: 10.7861/futurehosp.6-2-94.