Research in Education and Rehabilitation 2023; 6(2): 191-200. **DOI**: 10.51558/2744-1555.2023.6.2.191

CLASSIFICATION OF PATIENTS WITH VISUAL DISABILITY FOR THE BASIC FUNCTIONAL REHABILITATION PROGRAM

Casas Contreras Elladira [1*]

Faculty of Engineering and Basic Sciences of the Central University of Bogotá, Colombia

Original Scientific Article

Received: 17/09/2023
Accepted:05/11/2023

ABSTRACT

The rehabilitation needs of people with visual disabilities vary due to different factors: congenital or acquired eye diseases, new social interaction, adaptation to productive, academic, or similar life, among others, meaning a continuous readjustment of the treatment plan of the patient and increasing rehabilitation costs. The care history of the Rehabilitation Center for Blind Adults (CRAC) has followed the basic functional rehabilitation route of the Rehabilitation Manual of the Latin American Union of the Blind (FOAL-ULAC, 1999) and contains the variables that can efficiently classify these patients with the use of machine learning tools. The patient's demographic (gender, age group) and clinical data (visual condition, admission ophthalmological diagnosis, emotional state), along with the number of visits, were collected through non-probabilistic sampling; The multiple correspondence analysis resulted in inverse associations between the quantitative and categorical variables, there was also a positive correlation between the quantitative variables according to Pearson's coefficient. To define the classification target variable, the similar characteristics between the variables were grouped into two clusters, using the scikit-learn library in Python and the k-prototype algorithm; After having the objective-labeled variable, the supervised Decision Tree, Random Forest, Gradient Boosting and Logistic Regression models were trained and tested. These models gave an accuracy between 82% and 84%, the most effective being the Gradient Boosting model, whose class prediction was: of 516 True Positives, 11 are False Positives, of 94 True Negatives and 15 are False Negatives (recall 98%), and accuracy of 83%.

Keywords: Basic Function Rehabilitation, Supervised Machine Learning, Clustering, Predictive Models.

[1]**Correspondence to:**
Casas Contreras Elladira, Faculty of Engineering and Basic Sciences of the Central University of Bogotá
E-mail: ecasasc@ucentral.edu.co

Research in Education and Rehabilitation 2023; 6(2): 191-200. **DOI**: 10.51558/2744-1555.2023.6.2.191

INTRODUCTION

This project is focused on the classification of Basic Functional Rehabilitation patients using supervised machine learning. This type of rehabilitation is an essential part of the institutional rehabilitation model, taught by the Latin American Union of the Blind (ULAC) and the Once Latin America Foundation (FOAL). The project is developed in two research areas: Education and Health Sciences; It is a combined modality, because special aspects of education are addressed, such as the cognitive and manual processes of people with visual disabilities and health aspects such as occupational therapy and psychology care, all of this, under the patient's medical diagnosis.

Comprehensive rehabilitation from the basic functional to the inclusion of blind people in society is very important to increase their autonomy and independence and to reactivate them as a social being. The loss of the visual faculty decreases some functional and psychosocial aspects in some proportion depending on the age of the person who suffers from it, for example, cognitive development is affected by the decompensation that the person suffers from not having new visual experiences (Piaget, 1977); sociocultural relationships are diminished by the lack of interaction with other people, which in turn causes their social isolation (Vygotsky, 1934); The limited space-time orientation reduces their motor performance, causing them to even forget the existence of the other parts of the body (Merleau, 1945); and the loss of neuroplasticity because the consciousness to carry out movements based on the same neuronal activity is lost, since the brain is a closed system that permanently receives information from the external environment through the senses and the response function depends on this. information (Llinás, 2003), among other aspects that affect their development and therefore their daily lives. Although motor function involves models of brain function, today a novel concept from the field of neurorehabilitation known as Motor Learning is applied to achieve changes in people's bodily performance (Cano-de-la-Cuerda and Others, 2015); All these factors are key when formulating a good basic functional rehabilitation treatment plan. By classifying patients into an appropriate treatment plan, sufficient financial resources and professionals can be allocated, as the cost of treatments can be determined, followed up and monitored more effectively.

The classification methods used in Latin America are dictated by the World Health Organization (WHO) and the best known are: Patient Classification System (SCP) to determine the level of medical care based on clinical criteria (García, J. 2006), Disease Severity Index (ISE) to measure the severity of the disease and its impact on health (Medina, E. 2000), the International Classification of Diseases (ICD) used worldwide to standardize through codes the diseases in order to allow the systematic recording of mortality and morbidity in each country (WHO, 2022), the Model List of Essential Medicines (EML) used as a reference worldwide to guarantee access to safe and effective medicines, pharmaceutical policies focused on public health (WHO, 2021), the International Classification of Disability and Health Functioning to globally and systematically codify the restrictions or limitations of these people in information systems (WHO, 2001), the Electronic Medical Record adopted as a tool to improve the quality of medical care and other information platforms with the aim of characterizing the population with disabilities (Ministry of Health and Social Protection, 2023).

Research in Education and Rehabilitation 2023; 6(2): 191-200. **DOI**: 10.51558/2744-1555.2023.6.2.191

Data science today allows research in different areas of health, for example, in hospital care, studies were found to identify patients at risk of post-operative complications (Domínguez, 2023), studies to classify patients with complex diseases such as breast cancer (Agudelo, H. 2021), glaucoma (Espinosa, A and Others 2014), others for exploration of chronic diseases (see Gomez, D. 2020), classification of diagnostic aids such as the NEVO platform (Agudelo, H. 2021). Bello, J. 2020), diagnostic support analysis such as electroencephalograms using neural networks (Delgado, Karina, and Others 2019), among many more. Regarding research related to the classification of patients based on their visual capacity, no indexed articles were found, but there were systematic review studies: some to improve the visual rehabilitation of patients (Hernandez, J and Others 2022) and others for personalized visual rehabilitation training (Cardona, A. 2019). A portable device for assisted orientation and mobility was also found which uses virtual reality and artificial intelligence to adjust to the patient's needs (Solis, C. 2019). Data science currently expands the possibilities of analyzing own or open data from special rehabilitation entities, which combine interdisciplinary knowledge and are committed to innovation, taking advantage of available resources, which are in line with the development policies of each country.
Finally, and to contribute to technological advances in the field of comprehensive rehabilitation, this project aims to determine a novel method taking advantage of these techniques and tools of data analytics, to identify potential patients for basic functional rehabilitation from the first attention and classify them into the appropriate level of treatment.

MATERIALS AND METHODS

Sample of participants

This research uses convenience sampling given the availability of CRAC's own data, which is the result of comprehensive rehabilitation care in 3,217 people with visual disabilities.

Method of conducting research

Using the scientific method, we seek to resolve the research question: ¿How are Basic Functional Rehabilitation treatments related to the demographic, psychological or behavioral variables of the patient with visual impairment? analyzing whether these variables can be predicted using a classification model. The approach of this research combines discrete quantitative data and qualitative data, extracted from the CRAC databases using SQL statements; The measurements, data treatment and statistical analysis include numerical correlation, Pearson coefficient, multiple correspondence analysis, estimation graphs for categorical variables using the Python programming language libraries (pandas, seaborn, matplotlib). The characterization by clusters is achieved using the K-prototypes algorithm which finds similarities with the Euclidean distance for numerical variables and a dissimilarity measure for categorical variables. For all inferential analyzes the Scikit-learn library was used.

Research in Education and Rehabilitation 2023; 6(2): 191-200. **DOI**: 10.51558/2744-1555.2023.6.2.191

Measurement tools

The measurement of the characteristics of individuals is carried out using a form and diagnoses given by specialists. The data set used in the project is made up of individual records of people with visual disabilities who have the following attributes: gender, visual condition, age, emotional state, and admission diagnosis. Likewise, discrete numerical attributes are obtained for these people, quantifying care in the rehabilitation areas: abacus, braille, elements of communication, techniques of daily living, orientation and mobility, physical preparation, and cognitive and manual development.

Data procession methods

For data processing, the CRISP-DM SPSS Modeler methodology is used to obtain effective predictive models. Below are the phases executed:

Business understanding phase

Information is collected about the CRAC, its patients and the care processes based on documented information from the Integrated Management System; Data extraction was done from the databases, instantiated on a SQL Server v18.3.1 platform and Microsoft Windows Server 2019 Standard operating system on an HP Proliant DL380 Tenth Generation server. Files with *.xlsx and *.cvs formats were created.

Data understanding and preparation phase

The sufficiency and quality of the data for analysis is reviewed, performing actions such as: reduction of rows and columns, standardization, typographical corrections, creation of new attributes, elimination of duplicates, etc., using Python data processing and visualization tools.

Modeling phase

The behavior of the variables, their relationships or associations, the characteristics that allow groupings are studied, the importance of the variables for the different classification models is found, the classification algorithms are selected which support the types of project variables: Decision Tree, Logistic Regression, Random Forest and Gradient Boosting. The data set is partitioned like this: 70% for training and 30% for testing. In the modeling phase, the variables that best explain the target variable of the classification models, their pattern of behavior in prediction, performance and the degree of importance are identified. The best classification model is also selected.

Evaluation and distribution phase

It is defined that the selected model is technically correct and adequate, control and maintenance are planned according to its validity.

Research in Education and Rehabilitation 2023; 6(2): 191-200. **DOI**: 10.51558/2744-1555.2023.6.2.191

RESULTS AND DISCUSSION

The important categorical variables used in the study have a particular structure due to non-probabilistic sampling; the imbalance between their categories required regularization to avoid overfitting in the models.

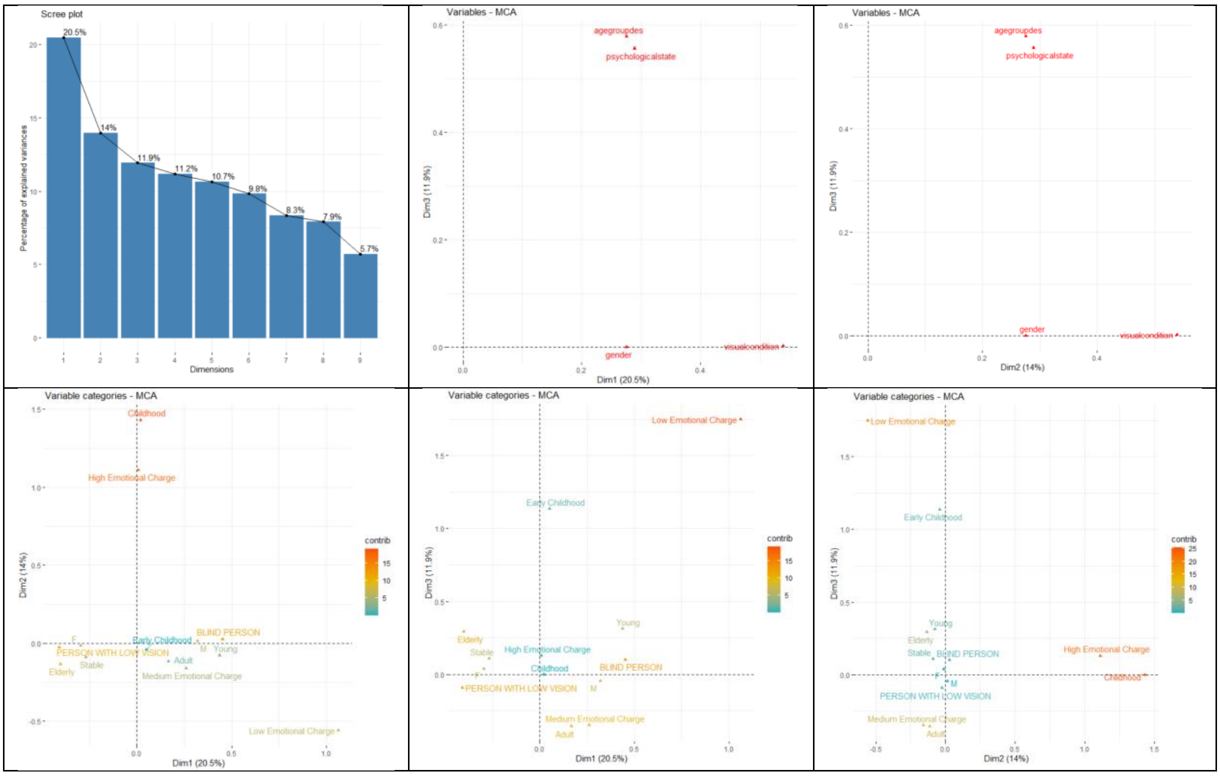

Figure 1. Multiple correspondence analysis of categorical variables

According to the Scree plot, 46% of the variance of the data is captured in the first three dimensions; Dimension 3 is characterized on the "y" axis with the variables age group and psychological status and on the "x" axis with the variables gender and visual condition, a strong association is observed between the categories of the "y" axis. A slight negative association is also observed between the categories: emotionally stable older adult women and young blind men.

Regarding the discrete variables, which quantify the care in each category of the issue variable, there is a minimum value of 0, this is because the people in the data set do not always receive all the categories in the treatment plan. Therefore, in the interpretation of quantitative analyzes the median metric will be essential.

The multivariate analysis did not clearly show a data pattern, but categories with a high correlation coefficient were obtained between the number of sessions of: physical preparation with abacus (0.80), physical preparation with orientation and mobility (0.72), orientation and mobility with abacus (0.66) and cognitive and manual development with physical training (0.62); the other categories had correlation coefficients below (0.6); The categories with a weak correlation coefficient were: cognitive and manual development with Braille (0.26).

Research in Education and Rehabilitation 2023; 6(2): 191-200. **DOI**: 10.51558/2744-1555.2023.6.2.191

The exploratory analysis of the data carried out so far responds to the objective of knowing the variables that contribute to the characterization of the treatment plans of Basic Functional Rehabilitation, however, among them there is no relevant variable for the classification objective. required; To find the target variable that groups the characteristics of the data set, it is decided to apply clustering techniques taking advantage of the criteria found for the relationship between the discrete variables and the association between the categorical variables: number of abacus sessions and physical training.

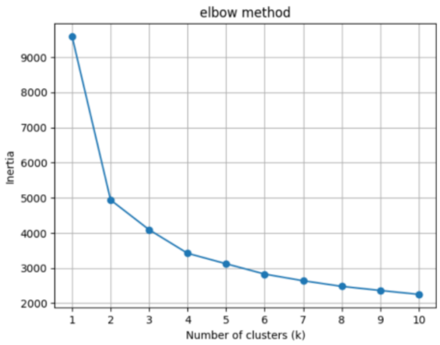

Figure 2. Optimal level of clusters

The optimal number of clusters found with the Elbow technique was k=2, the amount of variance explained in the retained dimensions had a balance of inertia close to 5000, the lowest obtained with respect to the other combinations. The clustering process with Skicit-learn's K-Prototypes algorithm allowed adding a column called "Cluster" to the data set, labeling each instance to one of the clusters 0 or 1, thus summarizing the relationship between the variables involved. ; The differences by cluster are specifically given by the psychological status variable while gender and visual status are very similar in their densities.

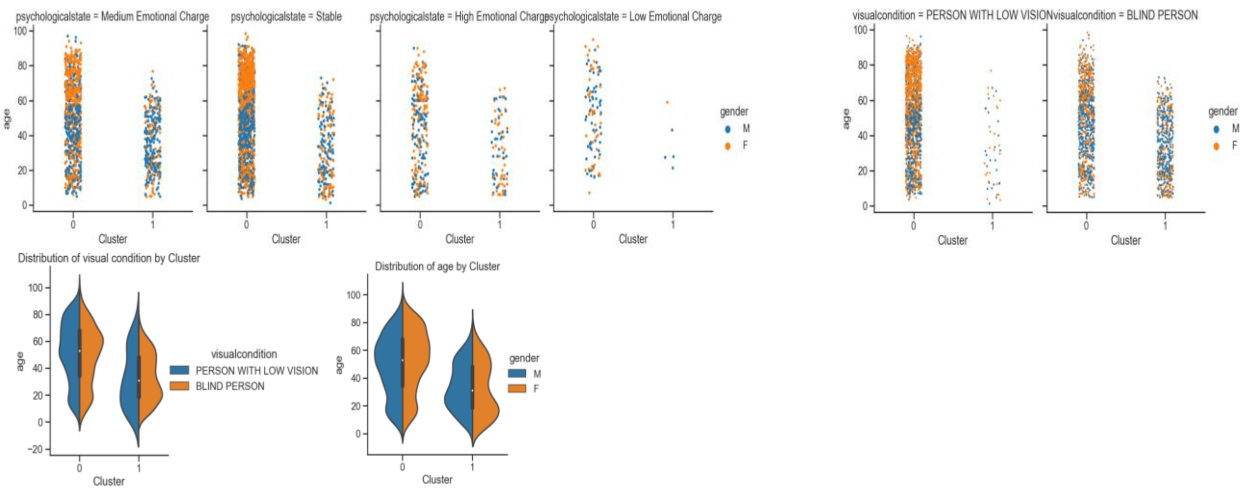

Figure 3. Distribution of characteristics by cluster

Once the target variable "Cluster" is defined, supervised machine learning algorithms are used to find the probability that it can or cannot belong to one of these clusters. The selected classification models handle both types of variables and data with multiple classes, in addition to improving the results by giving greater precision and avoiding overfitting. The training variables taken into account in the models are age group, psychological state, diagnostic code and the Cluster variable.

Research in Education and Rehabilitation 2023; 6(2): 191-200. **DOI**: 10.51558/2744-1555.2023.6.2.191

Table 1. Results of the classification models

Metrics	Logistic regression(LR)	Random forest(RF)	Gradient boost(GB)	Decision tree(DT)
accuracy	84.27%	82.07%	83.49%	82.70%
error	39.65%	42.33%	40.63%	41.59%
recall	97%	94%	98%	94%
f1-score	91%	90%	91%	90%
True positives	513	495	516	497
False positives	14	32	11	30
True negatives	86	82	94	80
False negatives	23	27	15	29
Probabilities in case classification	Class 0: 32.38% Class 1: 67.62%	Class 0: 16.65% Class 1: 83.35%	Class 0: 28.48% Class 1: 71.51%	Classified in Class 1

To compare the probability with which the models classify, a test was carried out with the record of an adult person admitted with High Emotional Burden and admission Diagnosis B24X (Human Immunodeficiency Virus Disease, without other specification). The algorithms used induce a single regression tree because the classification is binary (two classes), the best results are obtained from the Gradient Boosting model which “builds an additive model in advanced stages by optimizing arbitrary differentiable loss functions” (see Dangeti, 2017); To make a more reliable estimate of the performance of this model, cross-validation is performed, obtaining an average score of 0.80, which indicates good generalization capacity on unseen data; The variable that contributes the most to the prediction of the model is agegroupdes_Elderly with 32% importance, followed by other main variables that contribute in a lower percentage: psychologicalstate_Stable with 8%, disagnosticcode_H540 with 7%, diagnosticcode_H472 with 5%, agegroupdes_Adult with 4%, psychologicalstate_Low Emotional Charge with 3%; This ranking of variables is consistent with the multiple correspondence analysis in that older adults of any gender and visual condition, significantly influence the model, are those who probably require the greatest amount of attention for the Basic Functional Rehabilitation program; The classification with the admission diagnoses also allows us to identify emotionally stable women with low vision, with degeneration of the macula and the posterior pole of the eye, as the people who most frequently require increased attention. On the other hand, and less frequently, adult blind men with optic atrophy or diabetes and young blind men with optic atrophy also require increased amounts of care.

Although CRAC's own data have a structure that follows the classification guidelines stated in the introduction of this article, the training variables used for the model correspond to the specific categories of therapeutic care in the field of visual disability. For example, the admission diagnoses used in training the model are approximately 50% of the ICD-10 visual diagnoses, however, this patient classification method can be useful in any Latin American country, by preserving the structure of the rehabilitation model taught by the ULAC and FOAL organizations.

Research in Education and Rehabilitation 2023; 6(2): 191-200. **DOI**: 10.51558/2744-1555.2023.6.2.191

CONCLUSION

The analysis of the data from the cases treated for several years at the CRAC, provided the conditions for selecting the supervised Gradient Boosting classification model as the technically correct and appropriate model to assign the patient's treatment plan in the Basic Functional Rehabilitation program; Furthermore, the portability of the model makes it possible for the patient to detect this need from the first care. It is advisable to update the model at least every two years, because the training variables, especially those of the admission diagnosis, may or may not be renewed due to different changes in the factors that affect health.

REFERENCES

1. World Health Organization. (2022). Visual impairment and blindness. Recovered from https://www.who.int/es/news-room/fact-sheets/detail/blindness-and-visual-impairment
2. Hernández, Y. (2022). Postcensal Studies Characterization and changes in disability from the 2005 and 2018 censuses in Colombia (DANE – UNFPA). Chapter C. Aging, social protection and socioeconomic challenges: Characterization and changes in disability from the 2005 and 2018 censuses in Colombia: a look from the Social determinants. (Pages 26 - 74)
3. World Health Organization & World Bank. (2011). World disability report 2011. World Health Organization. https://apps.who.int/iris/handle/10665/75356
4. García, J. (2006). Complexity and function. Review of patient classification systems in geriatrics. Spanish Journal of Geriatrics and Gerontology. Vol 41 Number S1. (Pages 50 - 58). https://www.elsevier.es/es-revista-revista-espanola-geriatria-gerontologia-124-articulo-complejidad-funcion-revision-sistemas-clasificacion-13096078
5. Medina, E. Peratto, E. (2000). Validation of the disease severity index in inpatients in internal medicine wards of the teaching hospital. REV MED POST UNAH. Vol 5 Number 2. (Pages 133 – 137).
6. World Health Organization. (2022). Publication of the International Classification of Diseases ICD-11 2022. (ICD). Recovered from https://www.who.int/es/news/item/11-02-2022-icd-11-2022-release#:~:text=La%20Clasificaci%C3%B3n%20Internacional%20de%20Enfermedades,estandarizada%20en%20todo%20el%20mundo
7. World Health Organization. (2021). Model list of essential medicines. Recovered from https://www.paho.org/es/documentos/22a-lista-modelo-oms-medicamentos-esenciales-ingles
8. Ministry of Health and Social Protection. (April 17, 2023). This is how the electronic medical record will work in Colombia. Recovered from https://www.minsalud.gov.co/Paginas/Asi-funcionara-la-historia-clinica-electronica-en-Colombia.aspx
9. Piaget, J. (1977). The role of action in the development of thinking. In Knowledge and development (pp. 17–42). Springer US.
10. Vygotsky, L. (1934). Thought and Language. The role of Society, Culture and Language in the Development of the Human Being.

11. Merleau-Ponty, M. (1945). Phenomenology of perception. Translated Jem Cabanes. Barcelona; Mexico City; (Ed. Planeta-Agostini, 1997).
12. Llinás, R. (2001). The Brain and the Myth of the Self. Best Seller Best Health Book at the 2013 BookExpo America Book Fair.
13. Domínguez, E y otros (Septiembre de 2015). The Construction of the Gaze of Blind People. Orientation in Space. SEDICI Institutional Repository of the National University of La Plata. XXXIV Meeting and XIX Congress of ARQUISUR. Recovered from http://sedici.unlp.edu.ar/handle/10915/5141
14. Cano-de-la-Cuerda, A. et al (January-February 2015). Theories and models of motor control and learning. Clinical applications in neurorehabilitationTheories and control models and motor learning: clinical applications in neurorehabilitation. Sciencedirect. Neurology. Vol 30 issue 1. (Pages 32-41) Recovered from https://www.sciencedirect.com/science/article/pii/S0213485312000114
15. Baquet, J et al (1999). Technical Manual of Comprehensive Rehabilitation Services for Blind or Low Vision People in Latin America. https://www.foal.es/sites/default/files/docs/39_manual_de_rehabilitaci%26%23243%3Bn_0.pdf
16. Fernandez Llontop, R. J. (2021). USAT Thesis Repository Catholic University of Peru. Intelligent geolocating sensory cane to support the movement of blind people in the Regional Organization of the Blind of Peru–Chiclayo. Dspace repository. Recovered from https://tesis.usat.edu.pe/handle/20.500.12423/3213
17. Solís Yagual, C. R. (2018). Design of the prototype of a glove controlled with Arduino that allows the detection of obstacles through sensors to improve movement within the home for low-income people with visual disabilities (Doctoral dissertation, University of Guayaquil. Faculty of Mathematical and Physical Sciences. Networking and Telecommunications Engineering Career). Recovered from http://repositorio.unesum.edu.ec/handle/53000/1470
18. Hoyos Rivas, F. A. (2022). Development of a facial protector with ultrasonic technology and ir detection to improve the movement of blind people - Covid 19. Retrieved from http://repositorio.unac.edu.pe/handle/20.500.12952/6508
19. Espinosa Castañeda, A. P., Peña Martínez, V. & Carrizosa, M. (2014). Description and analysis of OCT, HRT and GDx diagnostic technologies in glaucoma: an approach for clinical optometry. Science & Technology for Visual and Eye Health, 12(2), 87-106. Recovered from https://ciencia.lasalle.edu.co/svo/vol12/iss2/9/
20. Bello, J. (2020). Rosario University Repository. “NEVO: Intelligent Electronic Medical Record for Visual Health”. Recovered from https://repository.urosario.edu.co/items/023706b3-31ef-4622-80b9-afe476f649fe
21. Mejía, R., Cubillos, L., & González, M. (2017). Electronic medical record in Latin America: current status and perspectives. Medical Journal of Chile, 145(1), 97-105. Recovered from https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0034-98872017000100014&lng=es&nrm=iso&tlng=es
22. Domínguez, S. Andrade, R. (2023). Artificial neural network to predict risk factors associated with postoperative complications secondary to pneumothorax treatment.

Research in Education and Rehabilitation 2023; 6(2): 191-200. **DOI**: 10.51558/2744-1555.2023.6.2.191

Colombian Journal of Surgery. Recovered from https://www.revistacirugia.org/index.php/cirugia/article/view/2225
23. Mera, L. Ochoa, J. (2021). Convolutional neural networks for independent component classification of rs-fMRI. ITM Magazine. Investigation article. Vol 24 No. 50. Recovered from http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0123-77992021000100097
24. Agudelo, H. Sariia, M. (2021). Breast cancer detection using histopathology images and pre-trained deep learning models. CESTA Journal of Computer and Electronic Sciences. Vol 2 No. 2. (Pages 27-36) Retrieved from https://revistascientificas.cuc.edu.co/CESTA/article/view/3922
25. Espinosa, A.y Otros (2014). Description and analysis of OCT, HRT and GDx diagnostic technologies in glaucoma: an approach for clinical optometry. Science & Technology for Visual and Eye Health, 12(2), 87-106. Recovered from https://ciencia.lasalle.edu.co/svo/vol12/iss2/9/
26. Alvarado, R. Llerena, J. (2021). Review of the literature on the use of Artificial Intelligence focused on the care of visual impairment. Scientific and Technological Magazine InGenio magazine of the Faculty of Engineering Sciences. Quevedo State Technical University, Ecuador. Vol5 No.1 Recovered from http://portal.amelica.org/ameli/jatsRepo/385/3852780002/index.html
27. Cardona, A. Vasquez, R. (2019). Mobility assistance devices in people with visual disabilities: A bibliographic review. Polytechnic Magazine, vol. 15, no. 28, pp. 107-116. Recovered from https://revistas.elpoli.edu.co/index.php/pol/article/view/1534
28. Fernandez, R. J. (2021). Intelligent geolocating sensory cane to support the movement of blind people in the Regional Organization of the Blind of Peru–Chiclayo.
29. Solísl, C. (2019). Design of the prototype of a glove controlled with Arduino that allows the detection of obstacles through sensors to improve movement within the home for low-income people with visual disabilities (Doctoral dissertation, University of Guayaquil. Faculty of Mathematical and Physical Sciences. Networking and Telecommunications Engineering Career).
30. Gómez, D. A., Hernández, J. A., Roa, L. M., & Delgado, O. (2020). Classification of patients with chronic diseases in Latin America using convolutional neural networks. Journal of Public Health, 22(1), pp. 39-45.
31. Ávila, J. et al. (2020). Artificial intelligence and its applications in medicine II. Primary Care. ScienceDirect. Vol 53 Issue 1. Pp. 81-88. Recovered from https://www.sciencedirect.com/science/article/pii/S0212656720301463
32. Hernández, J. et al (2022). Use of virtual reality and deep learning to improve visual rehabilitation in patients with homonymous hemianopia. Virtual Reality Magazine, 6(1), pp. 12-21.
33. Delgado,K.,Ledesma, S.,& Rostro, H.(2019) Electroencephalogram analysis using artificial neural networks. University Act 29, e1672. doi. http://doi.org/10.15174.au.2019.1672.
34. Dangeti, Pratap (2017). Statistics for Machine Learning : Build Machine Learning Models with a Sound Statistical Understanding. 1st ed. Birmingham: Packt Publishing, Limited, Chapters 3-4.