Support vector machine with optimized parameters for the classification of patients with COVID-19

Daniel Andrade-Girón [1], Edgardo Carreño-Cisneros [1], Cecilia Mejía-Dominguez [1], Julia Velásquez-Gamarra [1], William Marín-Rodriguez [1,2,*], Henry Villarreal-Torres [3], Rosana Meleán-Romero [4]

[1] Universidad Nacional José Faustino Sánchez Carrión. Huacho, Perú.
[2] Universidad Tecnológica del Perú. Lima, Perú.
[3] Universidad San Pedro. Chimbote, Perú.
[4] Universidad Cesar Vallejo. Lima, Perú.

Abstract

Introduction. The COVID-19 pandemic has had a significant impact worldwide, especially in health, where it is crucial to identify patients at high risk of clinical deterioration early.

Objective. This study aimed to design a model based on the support vector machine (SVM) algorithm, optimizing its parameters to classify patients with suspected COVID-19.

Methodology. One thousand patient records from two health establishments in Peru were used. After applying data preprocessing and variable engineering, the sample was reduced to 700 records. The construction of the model followed a machine learning methodology, using the linear, polynomial, sigmoid, and radial kernel functions, along with their estimated optimal parameters, to ensure the best performance.

Results. The results revealed that the SVM model with the linear and sigmoid kernels presented an accuracy of 95%, surpassing the polynomial kernel with 94% and the radial kernel (RBF) with 94%. In addition, a value of 0.92 was obtained for Cohen's kappa, which measures the degree of agreement between the predictions of the machine learning model and the actual results, which indicates an excellent deal for the linear and sigmoid kernel.

Conclusions. In conclusion, the SVM model with linear and sigmoid kernels could be a valuable tool for identifying patients at high risk of clinical deterioration in the context of the COVID-19 pandemic.

*Corresponding author. Email: wmarin@unjfsc.edu.pe, C25859@utp.edu.pe

1. INTRODUCTION

On December 21, 2019, doctors in China reported cases of atypical pneumonia in dozens of patients in the city of Wuhan (Mojica Crespo & Morales Crespo, 2020). Taking into account the accelerated growth of infections and deaths that occurred in early 2020, the Chinese authorities submitted a report to the World Health Organization. (Reyes Nuñez & Simón Domínguez, 2020). On March 11, 2020, the WHO, considering the spread dynamics and its great danger, declared the SARS-CoV-2 pandemic, mentioning that it was a highly lethal virus and virologically similar to SARS-Cov-1 (HcoV -229E) that appeared in 2009 and would change the course of human history (De León et al., 2020).

The SARS-CoV-2 pandemic has modified the development dynamics in different institutions worldwide (Zarei et al., 2022). In health, immediate attention is required in many activities, such as timely diagnosis and revitalizing early

care in health establishments, as a fundamental task of organizations (Rodríguez Yago et al., 2020). In health facilities worldwide, a comprehensive clinical spectrum has been observed in patients with COVID 19 and levels of severity ranging from an asymptomatic course to acute respiratory distress syndrome (ARDS) and even death (Aljameel, Khan, Aslam, Aljabri, & Alsulmi, 2021). This has generated a complex problem in the different health establishments that did not have the necessary machinery and sufficient personnel to respond to timely patient care objectively (Martínez Chamorro, Díez Tascón, Ibañez Sanz, Ossaba Vélez, & Borruel Nacenta, 2021). The lack of an early diagnosis and not acting on time in the treatment would increase the clinical risk of the patient with COVID-19. As a consequence, there is a high percentage of deaths in many countries of the world, including Peru (Mendoza-Ticona, and others, 2020) (Cura-González, Polentinos-Castro, Fontán-Vela, López-Rodríguez, & Martín-Fernández, 2022).

Timely detection of diseases is a priority in health systems because time is a fundamental resource in identifying clinical risk and patient treatment (Chávarry-Ysla et al., 2023; Gutiérrez Solano & Solorzano Bernita, 2022; Cortés, 2021; Llerena & Aucanshala Shigla, 2023). In medical emergency services, patients with COVID-19 should be classified according to the clinical range, as well as those at high risk and those with a poor prognosis (Balseca Valle & Noroña Salcedo, 2022; Kalezhi et al., 2022; Lasluisa-Toalombo et al., 2022). In many cases, the latter presents high comorbidity, multiple pathologies, and specific signs and symptoms (Rivera & del Pino Casado, 2020). It is necessary to have mechanisms that allow the classification of patients in real-time to optimize patient treatment time in a scenario of accelerated growth of those infected with SARS-CoV-2 (Mohammad, Aljabri, Aboulnour, Mirza, & Alshobaiki, 2022).

In many countries, it has been suggested to use respiratory triage and severity identification scales, as well as the risk of mortality in patients with suspected infection. The British Thoracic Society maintains that the National Early Warning Scale (NEWS) makes it possible to detect patients with SARS-CoV-2 infection in emergency services to decide admission or hospitalization, in addition to clinical judgment, such as the pneumonia severity index (PSI) or the CURB-65 scale (confusion, urea, respiratory rate, blood pressure and age $\geq$ 65 years) (Romero Hernández et al., 2020). It is necessary to apply efficient classification mechanisms in patients with COVID-19 who go through potentially complex and changing pathophysiological situations (Li, et al., 2021). Knowing the clinical condition of the patient and classifying it according to the clinical range is the first step for its treatment and stabilization, which should be done as early as possible (Lalueza, et al., 2015; Castellanos & Figueroa, 2023; Aljameel, Khan, Aslam, Aljabri, & Alsulmi, 2021).

In hospitals where there is a continuous increase in patients with COVID-19, the classification process must be automated (Aftab et al., 2022) to be a support system for the health professional and provide a timely response from classification (Dinar, et al., 2022; Olusegun Oyetola et al., 2023). Within the traditional alternatives, the statistical methods, minimum distance, and maximum likelihood of category are the statistical classifiers that depend on the multivariate normal distribution of the data of each class. If the normal distribution for each class is correct, then the classification has a minimum probability of generating error, and the maximum likelihood classifier is the appropriate choice (Kavzoglu & Colkesen, 2009; Ramírez et al., 2023). Indeed, the maximum likelihood method has limitations related to the assumptions of normal distribution and restrictions of the input data (Fuentes Marmolejo & Medina Parra, 2021; Simhan & Basupi, 2023).

An alternative is to design an automatic learning model using mathematical algorithms for classifying patients according to the clinical range with the Support Vector Machine (SVM) (Jain, Shankar, & Devi, 2020). This machine can train and learn (Sánchez Gómez, 2019). The importance of models based on machine learning lies in the possibility of making classifications and predictions with different models such as k-nearest neighbors, Bayes classifiers, decision trees, the support vector machine, etc. (Véliz Capuñay, 2020). In addressing non-linear and multivariate classification problems, the support vector machine efficiently solves classification and prediction problems (Guhathakurata, Kundu, Chakraborty, & Banerjee, 2021). Its success is due to its solid foundation of mathematical optimization; In addition, it has powerful tools and algorithms to find the solution in a non-linear context (Mosquera, Castrillón, & Parra, 2018).

The support vector machine is a new classification technique that has become relevant in recent years as it has been shown to perform better than traditional learning machines (Rizwan, Farhad, & Imam, 2021; Veliz Moreno et al. 2022; Tomás Portales et al., 2022), such as neural networks, for which it has been introduced as a powerful tool to solve classification problems (Betancourt, 2005; Macea-Anaya et al., 2023). A support vector machine aims to find an optimal hyperplane that separates one class from another, maximizing the distance between the points of different categories (Pisner & Schnyer, 2020). In an actual application of clarification or regression, the data is often not linearly separable, so it is necessary to design a non-linear support vector machine; this is possible by incorporating more polynomial characteristics (Ahmad, 2018; Campos Sánchez et al., 2022). Adding features is feasible to implement and works well with all machine learning algorithms. This method cannot handle complex data sets at a low polynomial degree. At high polynomial degrees, it creates many features, making the model too slow (Géron, 2020; Sebo et al., 2023). A viable alternative to solve combinatorial explosion is to apply SVM by adding polynomial features (Zohair et al., 2021). Another strategy, when a set of elements is not linearly separable, is to transform the original space utilizing a non-linear function into a Hilbert space (Díaz-Chieng et al., 2022; Campo León, 2017; Rincon Soto & Sanchez Leon, 2022).

The main objective of the kernel functions (i.e., linear, polynomial, radial basis, and sigmoidal) is to maximize the

margin between hyperplanes (Ahmad, 2018; Marinho de Sousa et al., 2022). The Hilbert Spaces with Reproductive Kernel theory shows that the kernel functions correspond to a dot product, which induces a linear space with greater dimension than the original space, possibly infinite (Aronszajn, 1944). This fact allows us to reproduce any linear algorithm in a Hilbert area or, equivalently, for any algorithm, there is a non-linear version. This fact is known as the kernel trick (Tharwat, 2019). From the investigations, an SVM-based model has not been obtained by optimizing the parameters and comparing kernel functions to classify patients with suspected COVID-19 accurately. Indeed, this research aims to design a model based on the SVM algorithm, optimizing its parameters, for the classification of patients with suspected COVID-19.

2. LITERATURE REVIEW

Significant advances have been made in classifying patients with suspected COVID-19, for example, by identifying chest X-ray images using the support vector machine. Kesav & Jubukumar (2022) mention that machine learning has advanced to solve a wide range of biomedical problems with high precision. The research uses a deep learning mechanism to identify chest X-ray images of patients with COVID-19. The Bayesian optimization technique found that the support vector machine is the most accurate among several compared classifiers. From the analysis of the research on the application of the support vector machine and the classification of patients with suspected COVID-19, it was observed that studies were identified in 52 Scopus articles, 11 IEEE Explore articles, and 17 Web of Science articles in classification by X-rays, comparison of models in predicting patients with COVID-19, and sentiment analysis on the management of the pandemic, among other studies divergent from the present investigation.

Dilmi (2022) has developed an approach for automatically diagnosing COVID-19 using chest X-ray images and AlexNet, VGG16, and VGG19 deep learning architectures to extract useful and relevant features. Then, he used as inputs a support vector machine with two discrete outputs: COVID-19 or No-findings. Furthermore, he used the Bayesian optimization (BO) algorithm to fit the parameters of the SVM classifier and choose the optimal parameters. The study's results indicate that the VGG16-SVM-BO and VGG19-SVM-BO models give the best performance with an accuracy of 99.47%.

Kesav & M.G. (2022) present an investigation that uses a deep learning mechanism to identify chest X-ray images of patients with COVID-19 and other patients with pneumonia in two- and three-class scenarios. The proposed approach employs the GoogLeNet architecture to extract features fed into different classifiers. With the Bayesian Optimization technique, the Kernel SVM is the most accurate among several compared classifiers. The model showed % overall accuracy of 98.31% for the two-class classification between COVID-19 and non-COVID-19 chest X-ray images and 98.60% for the Three-class classification problem between COVID-19, healthy, and viral pneumonia radiographs. The proposed system outperformed several existing architectures and was tested using smaller data sets to ensure robustness. Jeng & Hsieh (2021) used the SVM-supervised machine learning algorithm to build a model to analyze and predict the presence of COVID-19 in a person based on the symptoms experienced. Hyperparameters such as degree, cost, gamma, and kernels, including linear, radial, polynomial, and sigmoid, were tuned using R Studio to achieve the best possible model performance. The model was tested by ten-fold cross-validation, and the results show that the polynomial kernel with optimized hyperparameters produced the best accuracy of 98.02%.

The purpose of the researchers Singh et al. (2020) was to produce real-time forecasts using the SVM model. They investigated the Coronavirus disease 2019 (COVID-19) prediction of confirmed, deceased and recovered cases. The prediction will help plan resources, determine government policy, provide survivors with immunity passports, and use the same plasma for care. Data, including attributes such as confirmed location, deceased, COVID-19 recovery, longitude, and latitude, were collected from January 22, 2020, to April 25, 2020, worldwide. SVM was used to explore the impact on identification, deaths, and recovery. The research of (Zoabi, Deri, & Shomron, 2021; Ramírez Moncada et al., 2022) has been oriented towards the effective detection of SARS-CoV-2 that allows a rapid and efficient diagnosis of COVID-19 and can mitigate the burden on medical care, for which have developed prediction models that combine several characteristics to estimate the risk of infection. These models are intended to assist medical personnel worldwide in triaging patients, especially with limited healthcare resources. We established a machine learning approach that trained on the records of 51,831 people tested (of whom 4,769 were confirmed to have COVID-19). The model they developed predicted COVID-19 test results with high accuracy using only eight binary characteristics: sex, age ≥ 60 years, known contact with an infected individual, and the occurrence of five initial clinical symptoms.

After reviewing the literature, no studies were found that designed a machine learning model based on an automatic support machine that optimizes its parameters and compares the different kernels. Therefore, we aim to build a model based on the support vector machine algorithm to classify patients. The model implementation was developed in Python, a general-purpose programming language with a variety of packages for data science. (Valero Medina & Alzate Atehortúa, 2019; Andrade Gontijo et al., 2022; Santos Amaral et al., 2022).

3. MATERIALS AND METHODS

This research follows the positivist paradigm, quantitative approach, observational design without intervention, and predictive level. For the development of the machine learning model, the methodology for machine learning machine design was applied (Géron, 2020).

3.1. Data collection. One thousand records of patients diagnosed with SARS-CoV-2 infection admitted by the emergency service in health centers in Peru were collected. The variables considered were: age, gender, weight, height, respiratory rate, oxygen saturation, systolic blood pressure, heart rate, and temperature. Since the study is observational without intervention, informed consent was not requested, and the confidentiality of the data has been maintained.

3.2. Data preprocessing. Preprocessing is a fundamental phase that is performed before proceeding to the analysis. This ensures that the data is suitable for training and testing the model. In the present investigation, the data were analyzed with the Pandas package of the Python programming language. The data and records were validated concerning clinical risk to guarantee the adequate classification of the model; missing data were imputed; the characteristics were standardized; Fixed data imbalance issue. After analyzing, reducing the features, and balancing, a sample of 700 records was considered.

3.3. Selection of learning algorithm. There are various machine learning tools. The support vector machine was chosen for this study, which allows classification and prediction to optimize the different kernels and compare the results.

3.4. Implementation of the model and training. The program has been developed using the Python package scikit-learn. A training set, which is 70% of the sample, and a validation set, which is 30%, have been used, for which a program aimed at improving the results was designed. In this way, the error is minimized in each training, guaranteeing an incremental improvement in the model's efficiency.

3.5. Evaluation and validation. It has been verified that the model performs an efficient classification, for which the confusion matrix has been calculated. The performance of the model has been optimized from the linear, polynomial, radial, and sigmoidal kernels, adjusting the cross-validation by incorporating the Grid Search algorithm (Rios, Ulloa, & Borello Gianni, 2019), which allowed the C and Gamma parameters to be calculated automatically.

To compare the performance of the different kernels, sensitivity, specificity, and Cohen's kappa metrics have been used. We consider the number of false positives (FP) to the observations erroneously classified as correct. In the research, they are people who do not have COVID-19, and the model ranked them as COVID-19 (+) and the false negatives (FN) were the observations erroneously classified as accurate, that is, people who have COVID-19 and the model ranked them as COVID-19 (-). In addition, correctly classified results correspond to true positives (TV), people who have COVID-19 and were organized by the model as COVID-19 (+), as well as negatives (TV), people who do not have COVID-19 and the model classified them as COVID-19 (-). The four values form the confusion matrix, allowing validating of the model and choosing the optimal kernel (Moyano, Millán, Curiale, Moyano, & Portillo, 2018).

The metrics provided by the confusion matrix are accuracy, which represents the percentage of correct predictions compared to the total number of patients; precision, which refers to how close the result is to a prediction of the actual value; sensitivity (recall), which represents the rate of true positives (True Positive Rate) or TP, the proportion between the positive cases well classified by the model concerning the total number of positives and the specificity, which is the rate of true negatives (True Negative Rate) or TN, the proportion between the negative cases well classified by the model concerning the total number of negatives, Moyano et al. (2018).

4. RESULTS

During the data collection phase, 1,000 medical records of patients diagnosed with COVID-19 were identified in two health facilities in Peru. The data of patients who met the validation criteria about clinical risk were recorded. After performing the preprocessing and analysis of variables, which included reviewing missing values, selecting variables, and resolving the imbalance problem based on the clinical range, a sample of 700 records was considered.

The sample presented a mean age of 59 years and a standard deviation 15.90. In addition, it was found that 53.28% of the patients were male, and 63.71% were female. The sklearn package of the Python programming language was used to implement the model. The model's training was carried out using 70% of the random samples from the clinical histories, while the remaining 30% were used to carry out the test.

Below is a table evaluating the training results for the different kernels used: linear, polynomial, radial, and sigmoid.

Kernel	Clinic range	precision	recall	f1-score	support	Accuracy
Linear	Leve	0.93	1.00	0.96	67	
	Moderado	0.95	0.93	0.94	86	
	Severo	0.96	0.91	0.94	57	0.95
Poly	Leve	0.94	0.94	0.94	67	
	Moderado	0.91	0.94	0.93	86	
	Severo	0.98	0.93	0.95	57	0.94
Sigmoide	Leve	0.94	1.00	0.97	67	
	Moderado	0.94	0.93	0.94	86	
	Severo	0.96	0.91	0.94	57	0.95
Rbf	Leve	0.96	1.00	0.98	67	
	Moderado	0.94	0.92	0.93	86	
	Severo	0.93	0.91	0.92	57	0.94

Table 1. Evaluation results of the MSV models for each kernel.

The evaluation results reveal that the support vector machine model with linear kernel obtained high precision. Overall, he achieved an overall hit rate of 95%. When predicting patients with a severe clinical range, it showed an accuracy of 96% and a sensitivity (recall) of 91%. For patients with a moderate clinical range, it achieved 95% accuracy and 93% sensitivity. For patients with a mild clinical range, it demonstrated 93% accuracy and 100% sensitivity.

On the other hand, when using the polynomial kernel, the support vector machine model obtained a total success rate of 94%. When predicting patients with a severe clinical range, it presented an accuracy of 98% and a sensitivity of 93%. For patients with a moderate clinical range, it achieved an accuracy of 91% and a sensitivity of 94%. As for patients with a mild clinical range, it exhibited an accuracy of 94% and a sensitivity of 94%.

The results of the evaluation show that, when using the sigmoid kernel, the support vector machine model obtained a high total success rate of 95%. When predicting patients with a severe clinical range, it presented an accuracy of 96% and a sensitivity (recall) of 91%. For patients with a moderate clinical range, it achieved an accuracy of 94% and a sensitivity of 93%. As for patients with a mild clinical

range, it exhibited an accuracy of 94% and a sensitivity of 100%. In the case of the radial kernel (Rbf), the model achieved a total hit rate of 94%. When predicting patients with a severe clinical range, it showed an accuracy of 93% and a sensitivity of 91%. For patients with a moderate clinical range, it achieved an accuracy of 94% and a sensitivity of 92%. As for patients with a mild clinical range, it exhibited an accuracy of 96% and a sensitivity of 100%.

To guarantee the external validity of the classifying model and avoid its dependence on the number of characteristics of the data, different values of the C parameters (0.1; 1; 10; 50;100;500; 1000,10000) were explored. and gamma (1, 0.1, 0.01, 0.001, 0.0001) using an optimization process. The cross-validation technique was applied to the training matrix, and the Grid Search algorithm was used to obtain the best parameters for each kernel. In order to make an adequate comparison, the model's accuracy was evaluated, and Cohen's Kappa index was calculated to measure the agreement between the categorizations obtained and the reference labels. These measures make it possible to assess the accuracy and consistency of the model in its ability to classify cases correctly.

Table 2. Hyperparameters, accuracy, and Cohen's kappa for the evaluated kernels.

Algorithm	C	Gamma	Accuracy	Kappa	IC-95%
Linear	1000	0.001	95 %	0.92	0.90-0.94
Poly	10000	0.01	94 %	0.90	0.88-0.92
Sigmoide	10000	0.001	95 %	0.92	0.90-0.94
rbf	10	0.001	94 %	0.90	0.88-0.92

A comparison was made using the accuracy metric, representing the percentage of cases the model classified correctly. The results showed that both the linear kernel and the sigmoid kernel obtained an accuracy of 95%, which is higher than the accuracy of the polynomial kernel with 94%, and the radial kernel (RBF) also with 94%. In addition, Cohen's kappa index was calculated, which measures the degree of agreement between the predictions made by the machine learning model and the actual results. A value of 0.92 was obtained for the linear and sigmoid kernel, which indicates an excellent agreement between the model predictions and the actual results.

5. DISCUSSION

The early detection of time-dependent diseases, particularly COVID-19, is currently a priority in health systems (Martín-Rodríguez, 2020; Borges, 2023; Specht González et al., 2022). In a context where time is a decisive resource for predicting the evolution of the patient, different early warning scales have been developed to define the strategy to follow, depending on

the patient's diagnosis. These scales determine the most appropriate actions early (Garcia_Lamberechts et al., 2018; Tiwari et al., 2023). Among the many proposals that have been applied objectively in some health centers, the NEWS 2 and CURB 65 scales, taking into account the responses to the CURB 65 and CRB 65 scale (Gómez-Gómez A, 2020; Pérez Martínez et al., 2022; Gutiérrez et al., 2020; Vargas Chanes & Nieto Rivera, 2023), allow classifying patients to distinguish between those that require immediate medical attention and those that can be deferred with calculated risk.

In the context of accelerated growth of patients with COVID-19, using these scales in real-time is complicated, even more so if you do not have the necessary experience. The model based on a support vector machine will significantly help in the triage or emergency phase (Altman, 2005; Aljameel, Khan, Aslam, Aljabri, & Alsulmi, 2021; Oliva et al., 2022;). There are multiple models to solve the classification problem, for example, K-nearest neighbors, Bayes classifiers, decision trees, and neural networks. Each model presents its advantages and disadvantages (Véliz Capuñay, 2020). In the present study, it was decided to design a model based on a support vector machine because it has a solid mathematical foundation; In addition, it does not need all the available points to carry out the efficient

5

classification, which presents a fundamental advantage over other classification methods that use a high percentage of the training set to be able to find a separation between classes (Betancourt, 2005; Mesa Trujillo et al., 2022; Campo Leon, 2017).

The strategy of calculating the optimal separating hyperplane between the clinical range categories is a significant advantage over the neural network model since, in this model. The last solution will be satisfied with the first separation hyperplane found in the training process. (Granados, 2019; Gupta et al., 2023a; Gupta et al., 2023b; Santos Rocha & Ferreira Araújo, 2022). The model based on a support vector machine was implemented for the linear, polynomial, radial, and sigmoid kernels. The C and Gamma parameters were optimized using the Grid Search algorithm. The values of the parameters were C = 100 and Gamma = 0.001. In this way, significant results were obtained in each implemented kernel (Table 2). Given that the application of the model based on a support vector machine is carried out in the health system, we are interested in keeping false negatives lower since it would be detrimental for the patient to give a negative diagnosis of COVID-19 when in reality, we must opt for a higher sensitivity value, where false negatives are lower among the implemented kernels. We opted for the line and sigmoid kernel over the polynomial and radial kernels, as they have an accuracy of 95%, a higher sensitivity value and both a Cohen's kappa value of 0.92 higher than the linear and radial kernels.

5. CONCLUSIONS

The data for the proposed model were obtained from 1000 records of patients with suspected SARS-CoV-2 infection who were admitted by the emergency service in health centers in Peru. The variables considered were: age, gender, weight, height, respiratory rate, oxygen saturation, systolic blood pressure, heart rate, temperature, and diagnosis based on mild, moderate, and severe clinical risk. The data were analyzed with the Pandas package of the Python programming language. The data and records were validated concerning clinical risk to guarantee the adequate classification of the model; missing data were imputed; the characteristics were standardized; Fixed data imbalance issue. After analyzing, reducing the features, and balancing, a sample of 700 records was considered.

Different kernel functions have been used for performance analysis. The parameters for the kernel functions have been optimized: linear, radial, sigmoid, and polynomial. The comparison has been made with the accuracy (accuracy) that measures the percentage of cases that the model has succeeded in the classification. For the linear and sigmoid kernel, it is 95%, while for the polynomial and radial kernel, with an accuracy of 94%. Likewise, Cohen's kappa was calculated to carry out an adequate evaluation of the prediction, and the value 0.92

was obtained, which means the excellent concordance of the forecast through the linear and sigmoid kernel.

In contrast, the polynomial and radial kernel have obtained a value of 0.90. The rapid triage process for patients with COVID-19 reduces the number of clinical consequences and healthcare costs. The timely classification of patients with COVID-19 will allow better management of hospital surveillance, prevention, and control strategies.

Ethical aspects.

Acknowledgements.

Financing.

Conflicts of interest.
We declare that we have no conflicts of interest in this study.

References

[1] Aftab, M., Amin, R., Koundal, D., Aldabbas, H., Alouffi, B., & Iqbal, Z. (2022). Classification of COVID-19 and Influenza Patients Using Deep Learning. Contrast Media Mol Imaging. doi:10.1155/2022/8549707

[2] Ahmad, I. B. (2018). Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection. SPECIAL SECTION ON SURVIVABILITY STRATEGIES FOR EMERGING WIRELESS NETWORKS, 6(1). doi:10.1109/ACCESS.2018.2841987.

[3] Aljameel, S. S., Khan, I. U., Aslam, N., Aljabri, M., & Alsulmi, E. S. (2021). Machine learning-based model to predict the disease severity and outcome in COVID-19 patients. Scientific programming, 1-10. doi:https://doi.org/10.1155/2021/5587188

[4] Altman, D. G. (2005). Diagnostic (STARD) and prognostic (REMARK) studies. Medicina clinica, 125, 49-55, doi:10.1016/s0025-7753(05)72210-7. Medicina clinica,, 49-55. doi:10.1016/s0025-7753(05)72210-7

[5] Andrade Gontijo, M. C., Hamanaka, R. Y., & Ferreira de Araújo, R. (2022). Research data management: production and impact from Dimensions database data. Advanced Notes in Information Science, 2, 112–120. https://doi.org/10.47909/anis.978-9916-9760-3-6.89

[6] Aronszajn, N. (1944). La théorie générale des noyaux réproduisants et ses applications. . Proceedings of the Cambridge Philosophical Society,, 39, 133–153. doi: 10.1017/S0305004100017813.

[7] Balseca Valle, M., & Noroña Salcedo, D. (2022). Factores de riesgo e impacto psicológico en adolescentes de alta vulnerabilidad, durante confinamiento por COVID-19. Salud, Ciencia Y Tecnología, 2, 135. https://doi.org/10.56294/saludcyt2022135

[8] Betancourt, G. A. (2005). Las máquinas de soporte vectorial (SVMs). Scientia et technica., 1(27), 22.

[9] Borges, S. E. N. (2023). Produção de materiais educativos sobre a covid-19 e a Hanseníase para os povos indígenas Guajajara e Canela na Amazonia Legal Maranhense, Brasil. Salud, Ciencia Y Tecnología - Serie De Conferencias, 2(1), 116. https://doi.org/10.56294/sctconf2023116

[10] Campo León, E. (2017). Introducción a las máquinas de vector soporte (SVM) en aprendizaje supervisado. Trabajo de Fin de Grado en Matemáticas, Universidad de Zaragoza. Obtenido de https://zaguan.unizar.es/record/59156/files/TAZ-TFG-2016-2057.pdf

[11] Campos Sánchez, C. M., Guillén León, L. A., Acosta Yanes, R. C., & Gil Oloriz, M. A. (2022). Metaverse: the future of medicine in a virtual world. Metaverse Basic and Applied Research, 1, 4. https://doi.org/10.56294/mr20224

[12] Castellanos, S., & Figueroa, C. (2023). Cognitive accessibility in health care institutions. Pilot study and instrument proposal. Data & Metadata, 2, 22. https://doi.org/10.56294/dm202322

[13] Chávarry-Ysla, P. del R., Piscoya Angeles, P. N., Castagnola-Sánchez, C. G., Oliva-Yarlaqué, Y. M., Montalvo Chacón, N., Dubo Araya, P. A., & Campillay Campillay, M. . (2023). Métodos de gestión y duelo disfuncional ante el COVID-19. Salud, Ciencia Y Tecnología, 3, 345. https://doi.org/10.56294/saludcyt2023345

[14] Cortés, M. E. (2021). La pandemia de COVID-19: importancia de estar alerta ante las zoonosis. Revista de la Facultad de Medicina Humana, 21(1), 151-156. doi:http://dx.doi.org/10.25176/rfmh.v21i1.3451

[15] Cura-González, I., Polentinos-Castro, E., Fontán-Vela, M., López-Rodríguez, J. A., & Martín-Fernández, J. (2022). ¿Qué hemos dejado de atender por la COVID-19? Diagnósticos perdidos y seguimientos demorados. Informe SESPAS 2022. Gaceta Sanitaria, 36(1), s36-s43. Obtenido de https://doi.org/10.1016/j.gaceta.2022.03.003

[16] De León, J., Cruz, A. P., Ramírez, P. A., Valencia, Y. E., Carrillo, C. Q., & Ayala, E. V. (2020). SARS-CoV-2 y sistema inmune: una batalla de titanes. Horizonte médico, 20(2), 5. doi: https://doi.org/10.24265/horizmed.2020.v20n2.12

[17] Díaz-Chieng, L. Y., Auza-Santiváñez, J. C., & Robaina Castillo, J. I. (2022). The future of health in the metaverse. Metaverse Basic and Applied Research, 1, 1. https://doi.org/10.56294/mr20221

[18] Dilmi, S. (2022). Automatic COVID-19 diagnosis using deep learning features and support vector machines based on Bayesian optimization. ICAASE 2022 - 5th Edition of the International Conference on Advanced Aspects of Software Engineering, Proceedings. doi:10.1109/ICAASE56196.2022.9931584

[19] Dinar, A. M., Raheem, E. A., Abdulkareem, K. H., Mohammed, M. A., Oleiwie, M. G., Zayr, F. H., . . . Al-Andoli, N. M. (2022). Towards Automated Multiclass Severity Prediction Approach for COVID-19 Infections Based on Combinations of Clinical Data. Mobile Information Systems. doi: https://doi.org/10.1155/2022/7675925

[20] Elguera Chavarria, P., Prado Bush, O., & Barradas Ambriz, J. (2019). Implementación de una escala de gravedad para la activación del equipo de respuesta rápida: NEWS 2. Med Crit.33(2):98–103, url: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-8. NEWS 2. Med, 33(2), 98–103. Obtenido de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-8

[21] Fuentes Marmolejo, M. D., & Medina Parra, W. D. (2021). Diseño de un modelo predictivo-asistencial de pacientes infectados por Covid-19, mediante un modelo supervisado de Machine Learning basado en criterios de derivación hospitalaria o ambulatoria. Universidad de Guayaquil.

[22] Garcia_Lamberechts, E. J., Martín-Sánchez, F. J., Julián-Jiménez, A., Llopis, F., Martinez-Ortiz, M., Arranz-Nieto, M., . . . González-Del Castillo. (2018). Modelo de riesgo a 30 días en los pacientes ancianos con infección y sindrome de respuestas imflamatoria sistemática atendidos en los servicios de urgencias. Emergencias (Sant Vicente del Horts, 241-246.

[23] Géron, A. (2020). . Aprende machine learning con scikit-learn, keras y tensorflow. España: Anaya.

[24] Gómez-Gómez A, G.-E. M.-R. (2020). Diagnóstico y tratamiento temprano de neumonía ante la pandemia por COVID-19 en San Luis Potosí. ¿Es posible implementar una estrategia para lograrlo? NCT Neumol y Cirugía Tórax, 79(4), 214-220. doi:10.35366/97962

[25] Granados, M. E. (2019). Máquinas de soporte vectorial y árboles de clasificación para la detección de operaciones sospechosas de lavado de activos. Lámpsakos, 21, 26-38,. doi:doi: https://doi.org/10.21501/21454086.2904

[26] Guhathakurata, S., Kundu, S., Chakraborty, A., & Banerjee, J. S. (2021). A novel approach to predict COVID-19 using support vector machine. Data Science for COVID-19, 351 - 364. doi:10.1016/B978-0-12-824536-1.00014-

[27] Gupta, B. M., Kappi, M., Walke, R., & Bansal, M. (2023b). Covid-19 research in Bangladesh: A scientometric analysis during 2020-23. Iberoamerican Journal of Science Measurement and Communication, 3(1). https://doi.org/10.47909/ijsmc.445

[28] Gupta, B., Kappi, M. M., Walke, R., Bansal, M., & Mandal, A. (2023a). Covid-19 associated coagulopathy (CAC): Global research output, 2020-2022 . Iberoamerican Journal of Science Measurement and Communication, 3(2). https://doi.org/10.47909/ijsmc.48

[29] Gutiérrez Solano, J. G., & Solorzano Bernita, R. E. (2022). Complicaciones neuropsiquiátricas por COVID-19. Salud, Ciencia Y Tecnología, 2(S1), 223. https://doi.org/10.56294/saludcyt2022223

[30] Gutiérrez, E., Meller, L., Virdis, J. M., De Simón, F., Gurovich, C., & Fernández Leyes, L. (2020). Retweet or reply? Covid-19 and Twitter. The case of the city of Bahía Blanca (Argentina). AWARI, 1(2), e18. https://doi.org/10.47909/awari.79

[31] Jain, D., Shankar, V. G., & Devi, B. (2020). A Robust Approach of COVID-19 Indian Data Analysis Using Support Vector Machine. Lecture Notes in Electrical Engineering, 837, 355 - 366. doi:10.1007/978-981-16-8546-0_29

[32] Jeng, J.-H., Hsieh, J.-G., & Nayvé Villavicencio, C. (2021). Apoyo Vector MáquinaModelado para la predicción de COVID-19 basado en síntomas utilizando el lenguaje de programación R. ACM International Conference Proceeding Series, 65 - 70. doi:10.1145/3490725.3490735

[33] Kalezhi, J., Chibuluma, M., Chembe, C., Chama, V., Lungo, F., & Kunda, D. (2022). Modelling Covid-19 infections in Zambia using data mining techniques. Results in Engineering, 13(100363). doi:https://doi.org/10.1016/j.rineng.2022.100363

[34] Kavzoglu, T., & Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. International Journal of Applied Earth Observation and Geoinformation, 11(5), 352-359. Obtenido de https://doi.org/10.1016/j.jag.2009.06.002

[35] Kesav, N., & Jubukumar, M. G. (2022). A deep learning approach with Bayesian optimized Kernel support vector machine for Covid-19 diagnosis. Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization. doi:10.1080/21681163.2022.2099299

[36] Lalueza, A., Lora-Tamayo, J., de la Calle, C., Sayas-Catalán, J., Arrieta, E., Maestro, G., . . . Lumbreras, C. (2015). Utilidad de las escalas de sepsis para predecir el fallo respiratorio y la muerte en pacientes con COVID-19 fuera de Unidades de Cuidados Intensivos. Revista Clinica Española, s358-s363. doi:10.1016/j.rce.2020.10.004

[37] Lasluisa-Toalombo, P. M., Rosero Freire, D. A., Jacome-Lara, A. C., & Salazar-Garcés, L. F. (2022). Serological markers associated with poor prognosis in positive Covid-19 patients. Salud, Ciencia Y Tecnología, 2, 141. https://doi.org/10.56294/saludcyt2022141

[38] Li, Z., Wang, L., Huang, L. S., Zhang, M., Cai, X., Xu, F., . . . Liu, G. (2021). Efficient management strategy of COVID-19 patients based on cluster analysis and clinical decision tree classification. Scientific reports, 9626. doi:https://doi.org/10.1038/s41598-021-89187-3

[39] Llerena, C. A., & Aucanshala Shigla, E. V. (2023). Efectos adversos post vacunación contra el COVID-19 en adolescentes. Salud, Ciencia Y Tecnología, 3, 372. https://doi.org/10.56294/saludcyt2023372

[40] Macea-Anaya, M., Baena-Navarro, R., Carriazo-Regino, Y., Alvarez-Castillo, J., & Contreras-Florez, J. (2023). Designing a Framework for the Appropriation of Information Technologies in University Teachers: A Four-Phase Approach. Data & Metadata, 2, 53. https://doi.org/10.56294/dm202353

[41] Marinho de Sousa, R. P., & Shintaku, M. (2022). Data privacy policy: relevant observations for its implementation. Advanced Notes in Information Science, 2, 82–91. https://doi.org/10.47909/anis.978-9916-9760-3-6.112

[42] Martínez Chamorro, E., Díez Tascón, A., Ibañez Sanz, L., Ossaba Vélez, S., & Borruel Nacenta, S. (2021). Diagnóstico radiológico del paciente con COVID-19. Radiología (Madr., Ed. impr.). doi:https://doi.org/10.1016/j.rx.2020.11.001.

[43] Martín-Rodríguez, F. e. (2020). "¿Puede la Puntuación 2 de alerta temprana nacional prehospitalaria identificar a los pacientes en riesgo de mortalidad temprana intrahospitalaria? Un estudio de cohorte prospectivo multicéntrico". Corazón y pulmón. 585-591.

[44] Mendoza-Ticona, A., Valencia, M. G., Quintana, A. A., Cerpa, C. B., Garcia, L. G., Álvarez, C. C., & Rivero, V. J. (2020). Clasificación clínica y tratamiento temprano de la COVID-19. Reporte de casos del Hospital de Emergencias Villa El Salvador, Lima-Perú. Acta Médica Peruana, 37(2). Obtenido de http://dx.doi.org/10.35663/amp.2020.372.968

[45] Mesa Trujillo D, Zayas Argos CC, Verona Izquierdo AI, García Mesa I, López Zamora A. (2023). Caracterización de la capacidad funcional en Adultos Mayores.

Interdisciplinary Rehabilitation / Rehabilitación Interdisciplinaria, 1.

[46] Mohammad, M. A., Aljabri, M., Aboulnour, M., Mirza, S., & Alshobaiki, A. (2022). Classifying the mortality of people with underlying health conditions affected by COVID-19 using machine learning techniques. Applied Computational Intelligence and Soft Computing. doi: https://doi.org/10.1155/2022/3783058

[47] Mojica Crespo, R., & Morales Crespo, M. M. (2020). Pandemia COVID-19, la nueva emergencia sanitaria de preocupación internacional: una revisión. Medicina de Familia., 46, 65-77. doi:https://doi.org/10.1016/j.semerg.2020.05.010.

[48] Mosquera, R., Castrillón , O., & Parra, L. (2018). Máquinas de Soporte Vectorial, Clasificador Naïve Bayes y Algoritmos Genéticos para la Predicción de Riesgos Psicosociales en Docentes de Colegios Públicos Colombianos. Inf Tecnol, 29(6), 153–62. doi:10.4067/S0718-07642018000600153

[49] Moyano, L. G., Millán, D., Curiale, A., Moyano, N. A., & Portillo, J. M. (2018). Clinical risk factors data classification with machine learning methos. CLIAP-San Rafael Mendoza Argentina.

[50] Oliva, M., Silva Sandes, E., & Romero, S. (2022). Application of social network analysis to the institutional relations of the Higher Education System in the Rivera region-Livramento. AWARI, 3. https://doi.org/10.47909/awari.157

[51] Olusegun Oyetola, S., Oladokun, B. D., Ezinne Maxwell, C., & Obotu Akor, S. (2023). Artificial intelligence in the library: Gauging the potential application and implications for contemporary library services in Nigeria. Data & Metadata, 2, 36. https://doi.org/10.56294/dm202336

[52] Pérez Martínez, N. G., Vega Esparza, R. M., & López López, Y. Y. G. (2022). Analysis of organizational resilience in times of Covid-19: A bibliometric overview. Iberoamerican Journal of Science Measurement and Communication, 2(2). https://doi.org/10.47909/ijsmc.167

[53] Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In Machine learning. Academic Press. Obtenido de https://doi.org/10.1016/B978-0-12-815739-8.00006-7

[54] Ramírez, M. E., Ron, M., Mago, G., Hernandez–Runque, E., Martínez, M. D. C., & Escalona, E. (2023). Proposal for an epidemiological surveillance program for the prevention of occupational accidents and diseases in workers exposed to carbon dioxide (CO2) at a Venezuelan brewing company. Data & Metadata, 2, 55. https://doi.org/10.56294/dm202355

[55] Ramírez Moncada, J. A., Rodríguez Torres, E., & Zamora Reyes, J. R. (2023). Estrategias recreativas para suplir las carencias de niños y jóvenes en situaciones de la Covid-19 en el municipio Morón (Cuba). Región Científica, 2(1), 202328. https://doi.org/10.58763/rc202328

[56] Reyes Nuñez, M. A., & Simón Dominguez, J. I. (2020). Cómo estimar la letalidad del COVID-19. Revista Mexicana de Patología Clínica y Medicina de Laboratorio, 67(1), 4-8. doi: 10.35366/93845.

[57] Rincon Soto, I. B., & Sanchez Leon, N. S. (2022). How artificial intelligence will shape the future of metaverse. A qualitative perspective. Metaverse Basic and Applied Research, 1, 12. https://doi.org/10.56294/mr202212

[58] Rios, J., Ulloa, G., & Borello Gianni, A. (2019). Aplicación de regresión con vectores de soporte en un sistema recomendador de actividades sociales. XXV Congreso Argentino de Ciencias de la Computación

(CACIC)(Universidad Nacional de Río Cuarto, Córdoba. Códoba.

[59] Rivera, J. R., & del Pino Casado, R. (2020). Manual práctico de enfermería comunitaria. Elsevier.

[60] Rizwan, M. F., Farhad, R., & Imam, M. H. (2021). Support vector machine based stress detection system to manage COVID-19 pandemic related stress from ECG signal. AIUB Journal of Science and Engineering, 20(1), 8 - 16. doi:10.53799/AJSE.V20I1.112

[61] Rodríguez Yago, M., Alcalde, I., Gómez, R., Parias, M. N., Pérez, A., Canals, M., . . . Hernández-Tejedor, A. (2020). Recomendaciones sobre reanimación cardiopulmonar en pacientes con sospecha o infección confirmada por SARS_CoV-2 (COVID-19). Medicina Intensiva, 44(9), 566-576. doi:10.1016/j.medin.2020.05.004

[62] Romero Hernández, S., Saavedra Uribe, J., Zamarrón López, E., Pérez Nieto, O., Figueroa Uribe, A., Guerrero Gutierrez , M., & Lopez Ferminde, J. (2020). Protocolo de atención para COVID-19 (SARS-CoV-2) de la Sociedad Mexicana de Medicina de Emergencias. Sociedad Mexicana de Medicina de Emergencias.

[63] Sánchez Gómez, C. (2019). Desarrollo de soluciones software mediante aprendizaje automático en el ámbito de la salud. situación tecnológica y perspectivas. Obtenido de http://hdl.handle.net/10317/8110

[64] Santos Amaral, L., Medeiros de Araújo, G., & Reinaldo de Moraes, R. A. (2022). Analysis of the factors that influence the performance of an energy demand forecasting model. Advanced Notes in Information Science, 2, 92–102. https://doi.org/10.47909/anis.978-9916-9760-3-6.111

[65] Santos Rocha, E., & Ferreira Araújo, R. (2022). Rapid scientific communication in times of pandemic: the attention of pre-prints online about Covid-19. Advanced Notes in Information Science, 2, 103–111. https://doi.org/10.47909/anis.978-9916-9760-3-6.114

[66] Sebo, T. A. R., Oentarto, A. S. A., & Situmorang, D. D. B. (2023). "Counseling-Verse": A Survey of Young Adults from Faith-Based Educational Institution on the Implementation of Future Mental Health Services in the Metaverse. Metaverse Basic and Applied Research, 2, 42. https://doi.org/10.56294/mr202342

[67] Simhan, L., & Basupi, G. (2023). None Deep Learning Based Analysis of Student Aptitude for Programming at College Freshman Level. Data & Metadata, 2, 38. https://doi.org/10.56294/dm202338

[68] Singh, V., Poonia, R. C., Kumar, S., Dass, P., Agarwal, P., Agarwal, P., & Raja, L. (2020). Prediction of COVID-19 corona virus pandemic based on time series data using support vector machine. Journal of Discrete Mathematical Sciences and Cryptography, 23(8), 1583 - 1597. doi:10.1080/09720529.2020.1784535

[69] Specht González, A., Casas Núñez, Y. E., Meriño Pompa, Y., Naranjo Vázquez, S. Y., & Johnson Quiñones, M. (2022). Caracterización epidemiológica y complicaciones obstétricas de gestantes con diagnóstico de COVID-19. Salud, Ciencia Y Tecnología - Serie De Conferencias, 1(1), 6. https://doi.org/10.56294/saludcyt20226

[70] Tharwat, A. (2019). Parameter investigation of support vector machine classifier with kernel functions. Knowledge and Information Systems, 1269-1302. Obtenido de https://link.springer.com/article/10.1007/s10115-019-01335-4

[71] Tiwari, P., Chaudhary, S., Majhi, D., & Mukherjee, B. (2023). Comparing research trends through author-provided keywords with machine extracted terms: A ML algorithm approach using publications data on neurological disorders. Iberoamerican Journal of Science Measurement and Communication, 3(1). https://doi.org/10.47909/ijsmc.36

[72] Tomás Portales Y, Rodríguez Díaz O, Hernández Díaz ER, Álvarez Rosales Y, Dopico Álvarez M. (2023). Manifestaciones emocionales y apoyo social percibido en adultos mayores ante el impacto de la Covid-19. Interdisciplinary Rehabilitation / Rehabilitación Interdisciplinaria, 1.

[73] Valero Medina, J., & Alzate Atehortúa, B. (2019). Comparison of maximum likelihood, support vector machines, and random forest techniques in satellite images classification. . Tectura, 23(59), 13-26.

[74] Vargas Chanes, D., & Nieto Rivera, P. (2023). The importance of social ties in obtaining employment. AWARI, 3. https://doi.org/10.47909/awari.155

[75] Véliz Capuñay, C. (2020). Aprendizaje Automático: Introducción al Aprendizaje Automático. Fondo Editorial PUCP. Lima-Perú: Fondo Editorial PUCP.

[76] Veliz Moreno Z, Núñez Ravelo E, Arencibia Llanes L, Suarez Castillo IE, Contreras Tamargo Y. (2023). Ansiedad, depresión y estrategias de afrontamiento en pacientes convalecientes de Covid-19. Interdisciplinary Rehabilitation / Rehabilitación Interdisciplinaria, 1.

[77] Zarei, J., Jamshidnezhad, A., Haddadzadeh, S. M., Mohammad, H. A., Cheraghi, M., & Sheikhtaheri, A. (2022). Machine learning models to predict in-hospital mortality among inpatients with COVID-19: Underestimation and overestimation bias analysis in subgroup populations. Journal of Healthcare Engineering,. doi: https://doi.org/10.1155/2022/1644910

[78] Zoabi, Y., Deri, R. S., & Shomron, N. (2021). Machine learning-based prediction of COVID-19 diagnosis based on symptoms. Digital Medicine. doi:10.1038/s41746-020-00372-6

[79] Zohair, M., El-Sayed, A., Ashraf, E., Guesh, D., Osama, A. G., Abdallah, A. M., . . . Ibrahim, G. (2021). The COVID-19 pandemic: prediction study based on machine Learning models. ENVIRONMENTAL FACTORS AND THE EPIDEMICS OF COVID-19. doi:10.1007/s11356-021-13824-7