Classification of Acid-Base Disorders Through Machine Learning

Rodrigo Ruiz de la Peña Martinez[1] , Eutzhel André Del Val Saucedo[1] ,
Paola Aidee de la Cruz Gallardo[1] , Carlos Eduardo Cañedo Figueroa[1(✉)] ,
Abimael Guzmán Pando[2] , and Natalia Gabriela Sámano Lira[1]

[1] Facultad de Medicina y Ciencias Biomédicas, Universidad Autónoma de Chihuahua,
31125 Chihuahua, Mexico
ccanedo@uach.mx

[2] Laboratorio de Física Química Computacional, Facultad de Medicina y Ciencias Biomédicas,
Universidad Autónoma de Chihuahua, 31125 Chihuahua, Mexico

Abstract. The present research aims to develop a tool in the form of an algorithm that can provide an accurate diagnosis of a patient's acid-base balance without the need for manual calculations by the attending physician. During the research, three different algorithms (Bayesian, KNN and a neural network) were used, tested and compared in order to achieve a reliable result, speeding up the diagnostic process for the patient and reducing the human error that can arise from manual calculations. The results show that the Bayesian algorithm had the lowest performance achieved with an MF1 of 0.7850, followed by the KNN algorithm with an MF1 of 0.8553, the next was the neural network which obtained an MF1 of 0.9711. Finally, an algorithm assembled by the three mentioned above was tested generating an MF1 of 1, which was tested on 70 data samples. This suggests that the design can be used for the classification of acid-base problems.

Keywords: ANN · Acid-base diagnostic · Classification

1 Introduction

One of the inherent attributes to blood is its characteristic pH level, which is indicated in any solution by the pH scale. This scale ranges from 0 for strong acids to 14 for strong bases, with neutral value in the middle of both [1].

Normal pH values found in blood range from 7.35 to 7.45, typically staying around 7.40, which is considered slightly basic [1]. The measurement of these pH values is performed by the physician to assess the acid-base balance of a patient, and the values of carbon dioxide levels (CO_2) as well as blood bicarbonate levels (HCO_3) are used [2].

The acid-base balance mentioned earlier is constantly maintained by the precise action of both the renal system and the respiratory system, preserving the harmony between CO_2 and HCO_3-. Any deviation from the normal values would also imply an alteration in the functions of various organs in the body, leading to physiological disturbances and even death [1, 2]. The body also has buffering systems that can combine

with acids or bases, depending on the situation at hand, helping to prevent rapid changes in pH levels [2].

To date, the Henderson-Hasselbalch equation, in conjunction with various other equations, is used to express the patient's pH, which utilizes the values of carbon dioxide partial pressure (pCO2) and HCO3, determining the relationship between acids and bases. This equation aids in classifying different acid-base disorders that may occur in the body and helps define whether these problems are respiratory or metabolic in nature [2]. The Henderson-Hasselbalch equation is written as Eq. (1)

$$pH = 6.1 + log\left[\frac{\left(HCO_3^-\right)}{(0.03xpCO_2)}\right] \tag{1}$$

Currently, there are also acid-base maps available, especially the Du Bose basic map, which help visually identify the specific disorder [2]. The acid-base map is only helpful from a graphical perspective, but it's not practical for daily patient care use.

The problem that arises when using the Henderson-Hasselbalch equation and identification maps is that, when it comes to making a precise diagnosis, the physician is forced to manually perform these calculations or measurements once they have the patient's laboratory data. This process is unnecessarily time-consuming and can be tedious, leading to erroneous results. Therefore, there is now a search for the automation of this process, completely eliminating the possibility of human calculation errors. This automation is seen as a tool for healthcare personnel.

2 Methodology

2.1 Database

The information used in each of the developed algorithms was gathered from a synthetic database (SDB), which is artificially generated data rather than data based on real-world events [3]. This database was created by the collaborators of this work based on the acid-base map, which serves as a graphical tool for doctors to diagnose acid-base disorders. This enables them to visualize a graphical representation showing potential chemical states between blood concentrations of HCO3-, pH, and CO2.

The following acronyms were used to facilitate the later mention of the characteristic disorders, which correspond to the areas that compose the DuBose acid-base map. These classes are: normal state, *SN*, metabolic acidosis state, *SACM*, acute respiratory acidosis, *SACRA,* chronic respiratory acidosis, *SACRC*, metabolic alkalosis, *SAM,* acute respiratory alkalosis, *SARA,* and chronic respiratory alkalosis, *SARC*. Random points of the DuBose acid-base map distributed across these seven different regions were used to create the data for each of the possible patient conditions, resulting in a total of 268 data points *p(x,y)*, being *x* the HOC3 and *y* the pCO2 concentration respectively. From these points we randomly utilized 198 for training and 70 for testing considering the seven state classes. Table 1 shows the partitioning of the database data [2].

Table 1. Composition of the data set

| Full database	Data partitioning	Data consider by class	acronym
268 randomly selected points of the DuBose acid-base map	198 for training	17 -Normal state	SN
		33 -Metabolic acidosis	SACM
		33 -Acute respiratory acidosis	SACRA
		33 -Chronic respiratory acidosis	*SACRC*
		33 -Metabolic alkalosis	SAM
		33 -Acute respiratory alkalosis	SARA
		16 -Chronic respiratory alkalosis	SARC
	70 for testing	10 for each class	

2.2 Neural Network

The ANN models have the specific architecture format, which is inspired by a biological nervous system. ANN models are made up of neurons in a complex, nonlinear manner, just like the human brain. Weighted links are used to connect the neurons. [4]. In biological neural networks, learning is primarily driven by two forms of synaptic plasticity: long-term potentiation (LTP) and long-term depression (LTD). LTP strengthens the connections between neurons when they consistently exhibit correlated activity, while LTD weakens the connections when the activity is uncorrelated or weakly correlated. These changes in synaptic strength enable neurons to form new connections and modify existing ones, which is essential for learning and memory formation [5].

ANNs on the other hand are made up of an input node layer and an output node layer coupled by one or more hidden node layers. By activating functions, input layer nodes transmit information to hidden layer nodes, which then either activate or do nothing in response to the evidence. When the value of a certain node or collection of nodes in the hidden layer hits a certain threshold, a value is transmitted to one or more nodes in the output layer. The hidden layers apply weighting functions to the evidence. A lot of examples (data) must be used to train ANNs. [6].

For the realization of this project, a feedforward ANN with a single hidden layer consisting of 5 neurons was used, with the following hyperparameters: a learning rate of 0.1, 1000 epochs, and an error tolerance of 1e-29 these values were selected using a grid search technique, which shows the most optimal hyperparameters that will increase the efficiency of the algorithm. The value of the bias remains 1, but the values of their weights are being constantly updated. Figure 1 illustrate the ANN proposed model.

The output provided by the neural network determines the patient's acid-base balance situation as Eq. (2), Where: $Output_1$ represents a normal state, $Output_2$, chronic respiratory acidosis, $Output_3$, metabolic alkalosis, $Output_4$, acute respiratory alkalosis, $Output_5$, chronic respiratory alkalosis, $Output_4$, metabolic acidosis, and $Output_6$, acute
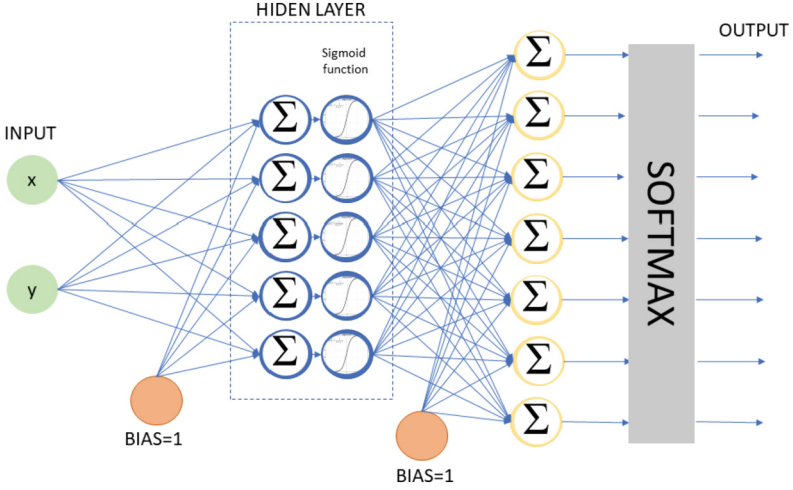
Fig. 1. Neural network

respiratory acidosis. Also, z_i stands for the final softmax output vector obtained after evaluating each i-th output, being N the total number of ANN outputs evaluated. Finally, arg max function is applied to obtain the winner class.

$$Z_i = \frac{e^{Output_i}}{\sum_{n \in N} e^{Output_n}}$$
$$Class = \arg\max(\{Z_i\}) \tag{2}$$

2.3 Bayesian Algorithm

The Bayesian algorithm (AB) functions as a probabilistic classifier that takes into account a particular feature and assigns a label to an element. AB operates using probability and employs empirical reasoning based on training data [7]. Its functionality lies in the correlation between a sample and its membership in a specific class, based on a set of features that are associated with different classes, each having its representative characteristics [8].

The AB training was performed using the set of classes $C = \{$SN, SACM, SACRA, SACRC, SAM, SARA, SARC$\}$ shown in Table 1. The probability of a training class P_C was obtained using Eq. (3), where T represents the total number of data points $p(x,y)$ consider for a c-th training class.

$$P_c = \frac{T_c}{\sum_{c \in C} T_c} \tag{3}$$

Also, BA involves Eqs. (4) and (5) that correspond to the mean and variance respectively.

$$\overline{p_c^l} = \frac{\sum_{j=1}^{T_c^l} p_{j,c}^l}{T_c^l}; l = \{x, y\} \tag{4}$$

$$\left(\sigma_c^l\right)^2 = \frac{\sum_{j=1}^{T_c^l} (p_{c,j}^l - \overline{p_c^l})^2}{T_c^l} \tag{5}$$

where j stands for the j-th point p evaluated on the characteristic l. Once the results of Eqs. (4) and (5) were obtained, Eq. (6) was used, where $P(c|p^l)$ correspond to the probability of a class c given a point p with features l [8, 9].

$$P\left(c|p^l\right) = \frac{1}{\sqrt{2\pi \left(\sigma_c^l\right)^2}} e^{-\frac{(p^l - \overline{p_c^l})^2}{2\left(\sigma_c^l\right)^2}} \tag{6}$$

Subsequently, Eq. (7) was implemented, which determines the relationship between all the probabilities obtained from Eqs. (3) and (6) for each of the seven classes.

$$Pr_c = \left(\prod P\left(c|p^l\right)\right) \cdot (P_c) \tag{7}$$

Equation (8) was used to obtain the evidence, which is the sum of all the different results from Eq. (7). This allows for the calculation of the likelihood that a given sample belongs to a particular class, based on Eq. (9) [8].

$$Ev = \sum_{c \in C} Pr_c \tag{8}$$

$$Ps_c = \frac{Pr_c}{Ev} \tag{9}$$

Finally, the AB determines the class to which it belongs by considering the Eq. (10)

$$Class = argmax(Ps_c) \tag{10}$$

where *Class* can have values from 1 to 7 according to C.

2.4 KNN

The KNN algorithm, also known as the k-nearest neighbors' algorithm, is a machine learning algorithm. Similar to the previous two algorithms, it is supervised, as the data provided to it already has labeled output values, allowing for the creation of a model to predict new input data [10].

The prediction of new input data is done by calculating the distance between the sample data and the data for each of the seven classes used in the training set [8]. In this case 5 neighboors (K) were used to classify the test dataset. This is done using

Euclidian distance between two samples p^l and q^l depicted on Eq. (11), being D the total components of the sample p^l, in this case $D = 2$ for the two components $l = \{x,y\}$.

$$d_{p,q} = \sqrt{\sum_{l=1}^{D} |p^l - q^l|} \qquad (11)$$

3   Results

The results of the evaluation on each previously trained algorithm are presented. The metrics show the precision, recall, and F1 scores (calculated using Eq. 12) for each class. The MF1 score (calculated using Eq. 13) is also shown for the overall results of the analysis algorithm. For this evaluation, we used 70 test data (10 per class) shown on Table 1.

$$F1_c = 2 \cdot \frac{precision \cdot recall}{precision + recall} \qquad (12)$$

$$MF1 = \frac{\sum_{c=1}^{7} F1_c}{7} \qquad (13)$$

Table 2 shows the metrics of the ANN classifier, with an MF1 score of 0.97114 (97.114%). This value is affected by the precision and recall scores for class one and class six. Where the precision score for class one is 0.8 (80%) and the recall for class six is 0.83 (83%).

Table 2. Effectiveness metrics for ANN

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Precision	0.8	1	1	1	1	1	1
Recall	1	1	1	1	1	0.833	1
F1	0.8888	1	1	1	1	0.9090	1
MF1	0.97114						

Table 3 shows the metrics for the KNN algorithm. The data shows a variation for each class, with the best-evaluated classes being class one and class six, both with a F1 score of 0.9523 (95.23%). However, this algorithm has an MF1 score that is 11.584% lower than the ANN algorithm.

Table 4 shows that the classes two and three were perfectly classified, with an F1 score of 1.0 (100%). However, the class six was not classified at all, with an F1 score of 0.0 (0%). This data affected the MF1 score, which was 0.785036 (78.5036%).

Table 3. Effectiveness metrics for KNN

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Precision	1	1	0.8	0.7	0.8	1	0.7
Recall	0.9090	0.7142	1	0.7777	0.8	0.9090	1
F1	0.9523	0.8333	0.8888	0.73684	0.8	0.9523	0.82352
MF1	0.8553						

The lower MF1 score is due to the fact that the class six was not classified at all. This means that the algorithm was not able to identify any of the instances of class six in the test dataset.

Table 4. Effectiveness metrics for Bayes

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Precision	1	1	1	1	1	0	1
Recall	0.83333	1	1	1	0.76923	1	0.625
F1	0.90909	1	1	0.94736	0.86956	0	0.7692
MF1	0.785036						

The neural network algorithm was the most efficient of the three individual algorithms, with an MF1 score of 97.11%. This suggests that the neural network can be used to classify the seven classes. However, there is an approximate 3% error rate in the classification of a data point. This problem could be fixed by retraining each algorithm with a larger dataset and comparing the future results with those obtained in this work.

4   Conclusion

At the end of the project, we managed to create a classification algorithm that, unlike the conventional method based on the Henderson-Hasselbach equation, uses only two parameters for classification. Additionally, our classification algorithm provided us with a 97% effectiveness score with its most reliable method, this being the neural network algorithm, which can serve as a prototype and later as a replacement for highly fallible and unpredictable human interference. This given that machine learning techniques provide numerous advantages over traditional diagnostic methods in the classification of acid-base disorders. These benefits encompass increased efficiency, enhanced accuracy, objectivity, scalability, decision support, early detection, and seamless integration with existing healthcare systems. The utilization of machine learning models enables the rapid processing of substantial amounts of data, facilitating prompt diagnosis in time-sensitive scenarios.

References

1. Lewis, J.L.: III Introducción al equilibrio ácido-básico. Manual MSD Versión Para Público General (2023). https://www.msdmanuals.com/es-mx/hogar/trastornos-hormonales-y-metab%C3%B3licos/equilibrio-%C3%A1cido-b%C3%A1sico/introducci%C3%B3n-al-equilibrio-%C3%A1cido-b%C3%A1sico
2. Del Pilar Triana Reyes, M., Pérez, V.H.E., Durán, D.P.J.: Chapter two - análisis de gases en la sangre. In: Estupiñan, V.H. (Ed.) Bases Para la Interpretación y Análisis de Gases Arteriovenosos. Cali, Colombia: Editorial Universidad Santiago de Cali; 2020, pp. 29–100 (2020). https://doi.org/10.35985/9789585583801
3. Durugkar, S.R., Raja, R., Nagwanshi, K.K., Kumar, S.: Introduction to data mining. Data Min. Mach. Learn. Appl. 1−19. https://doi.org/10.1002/9781119792529.ch1
4. Malekian, A., Chitsaz, N.: Concepts, procedures, and applications of artificial neural network models in streamflow forecasting. In: Advances in Streamflow Forecasting, pp. 115−147. Elsevier (2021). https://doi.org/10.1016/b978-0-12-820673-7.00003-2
5. Shehab, M., et al.: Chapter eight - artificial neural networks for engineering applications: a review. In: Elsheikh, A.H., Elaziz, M.E.A. (Eds.) Artificial Neural Networks for Renewable Energy Systems and Real World Applications, pp. 189–206. Academic Press (2022). ISBN 9780128207932. https://doi.org/10.1016/B978-0-12-820793-2.00003-3
6. Sadiq, R., Rodriguez, M.J., Mian, H.R.: Empirical Models to Predict Disinfection by-Products (DBPs) in Drinking Water: an Updated Review, pp. 324−338. Elsevier (2019). https://doi.org/10.1016/b978-0-12-409548-9.11193-5
7. Marimuthu, R., Shivappriya, S.N., Saroja, M.N.: Chapter 14 - a study of machine learning algorithms used for detecting cognitive disorders associated with dyslexia. In: Jude, H.D. (ed.) Handbook of Decision Support Systems for Neurological Disorders, pp. 245−262. Academic Press (2021). ISBN 9780128222713. https://doi.org/10.1016/B978-0-12-822271-3.00008-6
8. Figueroa, C.C., Chávez, H.G.: Diseño de algoritmo compuesto por machine learning y un modelo probabilístico para la detección de diabetes. In: Memorias del Congreso Nacional de Ingeniería Biomédica, vol. 8, No. 1, pp. 57−60 (2021). http://memoriascnib.mx/index.php/memorias/article/view/828/488
9. Chaitanya, B.K., Yadav, A., Pazoki, M., Abdelaziz, A.Y.: Chapter 8 - a comprehensive review of islanding detection methods. In: Zobaa, A.F., Abdel Aleem, S.H.E. (Eds.) Uncertainties in Modern Power Systems, pp. 211–256. Academic Press (2021). ISBN 9780128204917. https://doi.org/10.1016/B978-0-12-820491-7.00008-6
10. Team, D.: Descubra el algoritmo KNN : un algoritmo de aprendizaje supervisado. Formation Data Science | DataScientest.com (2022). https://datascientest.com/es/que-es-el-algoritmo-knn