

COMPARISON OF THREE ARTIFICIAL INTELLIGENCE ALGORITHMS FOR SEPSIS PREDICTION

Jelena Musulin¹, Daniel Štifanić¹, Ivan Lorencin¹, Sandi Baressi Šegota¹, Nikola Anđelić¹, Emanuel Borović², Alen Protić^{2,3}, Zlatan Car^{*1}

¹ University of Rijeka Faculty of Engineering, Department of automation and electronic, Rijeka, Croatia

² University of Rijeka Faculty of Medicine, Department of Anesthesiology, Reanimatology, Emergency and Intensive Care Medicine, Rijeka, Croatia

³ Clinical Hospital Center Rijeka, Clinic for Anesthesiology and Intensive Care, Rijeka, Croatia

*Autor za korespondenciju:

Zlatan Car

zlatan.car@riteh.hr

University of Rijeka Faculty of Engineering, Department of automation and electronic, Vukovarska 58, 51000 Rijeka, Croatia

SAŽETAK

Sepsa je najozbiljnija posljedica bakterijskih infekcija, drugim riječima, stanje opasno po život. Tijekom infekcije, reakcije nekih od komponenata urođenog imunološkog sustava mogu, pod određenim okolnostima, izazvati višestruku disfunkciju organa te oštećenje tkiva. Srećom, pravodobnim liječenjem moguće je spriječiti posljedice. Algoritmi koji se koriste za rješavanje problema predviđanja pojave sepse su: metoda potpornih vektora (SVM), k-najbližih susjeda (KNN) i umjetna neuronska mreža (ANN). Podaci korišteni u ovom istraživanju sastoje se od 23 medicinska parametra za svakog pacijenta, što je ukupno 2277 točaka u setu podataka. Za svaki od navedenih algoritama ispitane su različite kombinacije parametara, kao i različite vrste kernel funkcija (SVM), arhitektura (ANN) i broj najbližih susjeda (KNN). Eksperimentalni rezultati pokazali su da ANN metoda postiže najvišu AUC vrijednost (0.992) u usporedbi s drugim metodama poput SVM i KNN.

Ključne riječi: Sepsa, umjetne neuronske mreže, metoda potpornih vektora, k-najbližih susjeda

ABSTRACT

Sepsis is the most severe consequence of bacterial infections, in other words, a potentially life-threatening condition. During infection, some components of the innate immune response can, under certain circumstances,

cause multiple organ dysfunction and tissue damage. Fortunately, timely treatment can prevent the consequences. The algorithms used to solve the problem of predicting sepsis occurrence are Support-Vector Machine (SVM), K-Nearest Neighbors (KNN) and Artificial Neural Network (ANN). The dataset used for this research consists of 23 different medical parameters for each patient, which is the total of 2277 data-points. For each of the aforementioned algorithms, different combinations of parameters were examined as well as different types of kernel functions (SVM), architectures (ANN), and number of nearest neighbors (KNN). The experimental results showed that the ANN approach achieves the highest AUC value (0.992) compared to the other approaches like SVM and KNN.

Keywords: Sepsis, Artificial Neural Networks, Support Vector Machines, K-Nearest Neighbors

INTRODUCTION

Some machine learning models can use a large number of variables to predict outcomes of various diseases. Those methods can facilitate automation and application within a clinical decision support system, especially when dealing with severe cases such as sepsis (1). Sepsis is an inflammatory response of the body, often resulting in multiple organ dysfunction or tissue hypoperfusion, and even death (2). It manifests itself in three different stages: sepsis, severe sepsis and septic shock, with each of these stages

being more serious than previous one (3). In hospitals, mortality from sepsis is unacceptably high, and the best treatment strategy is still uncertain (4). It is important to predict the possible occurrence of sepsis while it is still in its early stages. For these reasons, the application of artificial intelligence (AI) in sepsis prediction is being imposed. Methods like Artificial Neural Network (ANN), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) are widely used in computation biology because of their high precision, ability to deal with high dimensions and large databases and their flexibility in modeling diverse data sources (5, 6). In addition to computational biology, these methods can be used in medicine (7) and other various fields of science, technology and maritime (8 – 10). Rapid interpretation of clinical data for the purpose of patient classification and predicting outcomes is paramount in emergency departments, and one of such emergencies is sepsis (11). In medicine, the cost of misclassifying one category is higher than in other fields, so it is preferable to use a classifier with selective sensitivity or specificity (12). In this research, the idea is to implement three different AI algorithms for sepsis prediction in order to compare classification performance. From the foregoing, the following hypotheses can be drawn:

- to achieve a high Area Under the ROC Curve (AUC), and
- to investigate the influence of different optimization algorithms (ANN), kernel functions (SVM) and number of nearest neighbors (KNN) on AUC value.

MATERIALS AND METHODS

This section is divided in Dataset Description and Methods Description. The Dataset description gives a short description of medical parameters used for sepsis diagnostic. The Methods Description provides a brief overview and mathematical description of ANN, SVM, and KNN methods used in this research as well as a subsection where the Performance Evaluation is described.

Dataset description

The dataset used in this research was obtained from the Clinical Hospital Center in Rijeka and consists of 23 different medical parameters for each patient. These parameters are labeled with “True” for having sepsis and “False” for not having sepsis. The parameters together with their marks and units of measurement are shown in Table 1.

The dataset contains information of medical parameters for 99 patients which is a total of 2277 data-points. 75% (1702 data-points, i.e. 74 patients) are used as a training set and 25% (575 data-points, i.e. 25 patients) as a testing set. The dataset is nearly balanced, with a total of 45% positive examples. Before implementation of the AI algorithms, the data needs to be preprocessed (13). Pre-processing techniques used in the research are: data normalization, and shuffling the order of the patients’ data (14). Afterwards, prepared data is used to solve the problem of predicting sepsis occurrence in patients.

Methods description

In this section, an overview of AI algorithms used for sepsis prediction is given, as well as description of the Performance Evaluation method.

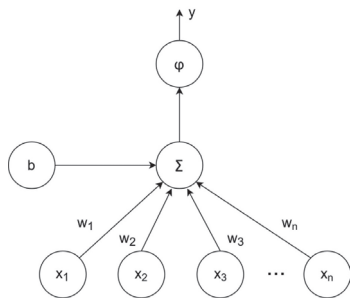


Fig. 1. Mathematical model of an artificial neuron

Table 1. Medical parameters used for sepsis diagnostic, their marks and units of measurement

Parameter	Mark	Unit
Thrombocytes	Tr	109/l
Leukocytes	L	109/l
Urea	U	mmol/l
Creatinine	Kr	μmol/l
Glucose	GUK	mmol/l
Sodium	Na	mmol/l
Potassium	K	mmol/l
Calcium	Ca	mmol/l
Bilirubin	BI Luk	μmol/l
Albumin	Alb	g/l
C-reactive protein	CRP	mg/l
Lactate	Lac (ak)	mmol/l
Procalcitonin	PCT	μg/l
pH	pH	-
Partial pressure (CO2)	pCO ₂	kPa
Partial pressure (O2)	pO ₂	kPa
Bicarbonate	HCO ₃	mmol/l
Total carbon dioxide	tCO ₂	mmol/l
Surplus amount of base	BE	mmol/l
Oxygen saturation	sO ₂	%
Oxygen saturation (vein blood)	sO ₂ (vk)	%
Prothrombin time	PV	-
International normalized ratio	INR	-
Sepsis	Y	True or False [0 or 1]

Artificial Neural Network method

Artificial neural networks (ANN’s) consist of artificial neurons grouped into three or more interconnected layers. These layers are categorized into three classes: input, hidden (may include more than one layer), and output (15). The way artificial neural networks store knowledge is represented by assigning and subsequently modifying the numerical values of weights. In addition to weights, a bias can be used, which can be considered as additional compensatory value for fine-tuning of parameters (16). Furthermore, the idea of the activation function is to decide, based on the results of mathematical operations of an artificial neuron, whether to activate the neuron and continue propagation through the network (17). For the purpose of this research, the ANN is trained using a supervised learning method (backpropa-

gation). The mathematical operations performed by an artificial neuron in a network are shown in Eq. (1)

$$y = \varphi \left[\sum_n (w_n x_n) + b \right]. \quad (1)$$

An equivalent mathematical model of an artificial neuron is shown in Fig. 1.

In this research, ANN is trained using three different optimization algorithms, and these are:

- Stochastic Gradient Descent (SGD),
- Adam, and
- Root Mean Square Propagation (RMSprop).

Support Vector Machine method

Support Vector Machine (SVM) is a classifi-

cation algorithm that finds a hyperplane of separation of different classification groups with the highest possible margin (18). The idea of SVM for linearly separable binary sets is to design a hyperplane that classifies all training vectors in two classes (19), as shown in Fig. 2.

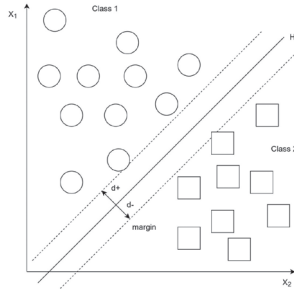


Fig. 2. Optimal hyperplane and maximum margin for given classes

Margin of Separation (d) is defined as the separation between the closest data-point and the hyperplane. The optimal hyperplane is equally distant from the nearest representatives of both classes and can be defined as

$$w^T x + b = 0, \quad (2)$$

where:

w is a weight vector,
 x is input vector, and
 b is bias.

The margin value can be defined as

$$\rho = \frac{2}{\|w\|}. \quad (3)$$

In order to maximize the margin, the method of Lagrange multipliers is used:

$$L = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N \alpha_i [y_i (w^T x_i + b) - 1]. \quad (4)$$

The aforementioned mathematical explanation of the SVM method has been based on the assumption that input data is linearly separable. If the problem is not linearly separable, the entire learning base needs to be transformed into another 3D space (20, 21). Furthermore, Φ represents the transformation operator.

Kernel function “K” can be defined as

$$K(x_i, x_j) = \Phi(x_i)\Phi(x_j). \quad (5)$$

In this research, the following three different kernel functions are used:

Linear,
 Radial Basis Function (RBF), and
 3rd degree Polynomial.

K-Nearest Neighbor method

K-Nearest Neighbor (KNN) algorithm is one of the most used supervised machine learning methods (22). It is a non-parametric algorithm that classifies new input data into one of the classification groups (23). The term non-parametric means that the algorithm makes no assumptions about the underlying distribution of the data. In other words, the structure of the model is determined directly from the learning data. There are several ways to determine the distance between vectors within multidimensional space (24). The most common and appropriate way for classification problems is to use the Euclidean distance formula, which can be defined as

$$d(X_i, X_j) = \sqrt{\sum_{m=1}^N (x_{im} - x_{jm})^2}. \quad (6)$$

The Eq. (6) is valid assuming that all parameters of input data have the same weight distribution. In other words, it is assumed that all parameters affect the final prediction equally. If the foregoing is not the case, it is necessary to find a mathematical expression that will scale the influence of an individual parameter on the final prediction. Such an expression can be defined as

$$d(X_i, X_j) = \sqrt{\sum_{m=1}^N w_m (x_{im} - x_{jm})^2}. \quad (7)$$

Eq. (7) shows that the weight coefficient w_n is added to the Euclidean distance formula (25). In this research, four different numbers of neighbors are used: 1, 3, 5, and 7.

Performance Evaluation

In this research, the primary metric for comparison is the Area Under the Receiver Operating Characteristic curve (AUC) (26). Higher AUC values indicate higher classification

accuracy. ROC curve can be plotted with false positive rate (FPR) on the x-axis and true positive rate (TPR) on y-axis (27). FPR can be defined as

$$FPR = 1 - Specificity, \quad (8)$$

and TPR can be defined as

$$TPR = Sensitivity. \quad (9)$$

RESULTS AND DISCUSSION

In this section, AUC values of various AI algorithms are compared. AUC values of various configurations for each method are also compared, such as optimization algorithms, kernel functions, and the number of nearest neighbors.

RESULTS

Using an ANN model based on trained routine, output values are obtained for each patient separately. The output determines whether the patient is in sepsis group or not. Not all parameters have equal weight or importance in the onset of sepsis, so it was necessary to adjust the parameters within the neural network to obtain the optimal solution (28). The ANN architecture that gives the optimal results is presented in Table 2.

Table 2. ANN architecture utilized for early detection of sepsis with a number of nodes per layer

Layer	Number of nodes	Activation function
Input Layer	23	/
Hidden Layer 1	8	ELU
Hidden Layer 2	32	ELU
Hidden Layer 3	16	ELU
Output Layer	2	softmax

Using the architecture shown in Table 1, three optimization algorithms were compared to obtain the highest AUC. The highest AUC value was achieved using SGD optimizer, as shown in Fig. 3.

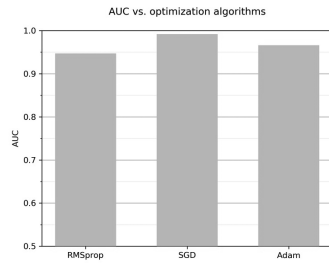


Fig. 3. AUC value versus the optimization algorithms for ANN designed with ELU activation function

Cross-Entropy was used as a loss function for measuring the performance of a classification model, and the highest AUC value achieved was 0.992.

SVM method has proven to be suitable for solving the problem of predicting sepsis occurrence in patients (29). If the patient is in the non-sepsis group, the algorithm will show “0”, and if the patient has sepsis or is in the high-risk group for sepsis, the algorithm will show “1”. The optimal results on the test set were given by the Linear kernel function with an AUC value of 0.958. AUC value for each kernel function is shown in Fig. 4.

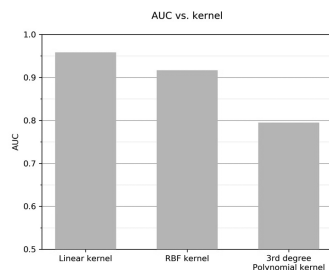


Fig. 4. AUC value versus the kernel functions for SVM

Since parameters strictly related to the onset of disease cannot be accurately determined in this research, and the initial distribution of the parameter is relatively unknown, the implementation of the KNN method has proved to be successful with reasonable AUC value (30, 31). The choice of variable “k” directly affects the final prediction, so the selection of that variable depends on the characteristics of the problem that the algorithm seeks to solve. Increasing the value of variable “k” reduces the impact of noise and irrelevant data in the input data (32). AUC value for different number of nearest neighbors is shown in Fig. 5.

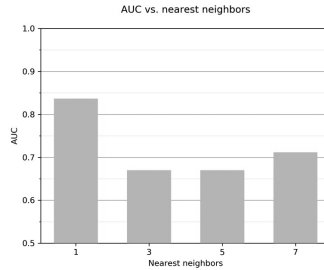


Fig. 5. AUC value versus the number of nearest neighbors for KNN

Value k = 1 was taken as the best solution, where the AUC value is highest (AUC = 0.837). The highest AUC values of three proposed methods are compared in Fig. 6.

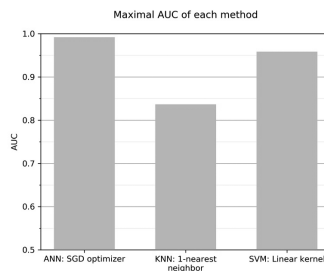


Fig. 6. AUC values for each configuration of three proposed methods

DISCUSSION

Based on all the results shown, maximal AUC was achieved using ANN method with SGD optimizer (learning rate of 0.01), Cross-Entropy as a loss function, and model architecture shown in Table 2. Furthermore, it can be noticed that the best results of three proposed methods had an AUC value of 0.837 or higher, as is shown in Table 3.

Table 3. Performance comparison of different methods for early detection of sepsis

Method	AUC values
ANN – SGD optimizer	0.992
ANN – Adam optimizer	0.966
SVM – Linear kernel	0.958
ANN – RMSprop optimizer	0.947
SVM – RBF kernel	0.917
KNN – 1 nearest neighbor	0.837

SVM – 3rd degree Polynomial kernel	0.795
KNN – 7 nearest neighbors	0.712
KNN – 3 nearest neighbors	0.670
KNN – 5 nearest neighbors	0.670

The ANN approach described in this research successfully predicted 24 out of 25 patients. According to research (33), the best performing method is ANN with 5 hidden layers, SGD optimizer with a learning rate of 0.01, and Mean Squared Error (MSE) as a loss function. Such an approach successfully predicted 15 out of 17 patients, where the total number of patients in dataset was 81.

CONCLUSION

In this research, three AI algorithms for predicting sepsis were successfully implemented. The highest AUC value achieved was 0.992, although the dataset was relatively small. The satisfactory result came from the choice of optimal activation functions, parameters and architecture configuration. All three algorithms have their advantages and disadvantages, however, all of them can be suitable for solving this type of problem. Implemented algorithms showed promising results on a relatively small dataset. With more data, these algorithms have real potential to be implemented and used in real-world scenarios. However, with existing dataset, implemented algorithms can be used in parallel with the existing diagnostic methods.

ACKNOWLEDGEMENTS

This research has been (partly) supported by the CEEPUS network CIII-HR-0108, European Regional Development Fund under the grant KK.01.1.1.01.0009 (DATA-CROSS), project CEKOM under the grant KK.01.2.2.03.0004 and University of Rijeka scientific grant uniri-tehnic-18-275-1447.

REFERENCES

1. Taylor RA, Pare JR, Venkatesh AK, Mowafi H, Melnick ER, Fleischman W, et al. Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data-driven, machine learning approach. *Academic emergency medicine*. 2016 Mar;23(3):269-78.
2. Stojkovic I, Obradovic Z. Predicting sepsis biomarker progression under therapy. In 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS) 2017 Jun 22 (pp. 19-24). IEEE.
3. Angus DC, Van der Poll T. Severe sepsis and septic shock. *N Engl J Med*. 2013 Aug 29;369:840-51.
4. Schorr CA, Zanutti S, Dellinger RP. Severe sepsis and septic shock: management and performance improvement. *Virulence*. 2014 Jan 1;5(1):190-9.
5. Vieira SM, Mendonça LF, Farinha GJ, Sousa JM. Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients. *Applied Soft Computing*. 2013 Aug 1;13(8):3494-504.
6. Lee BH, Scholz M. A comparative study: Prediction of constructed treatment wetland performance with k-nearest neighbors and neural networks. *Water, Air, and Soil Pollution*. 2006 Jul 1;174(1-4):279-301.
7. Lorencin I, Anđelić N, Španjol J, Car Z. Using multi-layer perceptron with Laplacian edge detector for bladder cancer diagnosis. *Artificial Intelligence in Medicine*. 2020 Jan 1;102:101746.
8. Lorencin I, Anđelić N, Mrzljak V, Car Z. Genetic Algorithm Approach to Design of Multi-Layer Perceptron for Combined Cycle Power Plant Electrical Power Output Estimation. *Energies*. 2019 Jan;12(22):4352.
9. Baressi Šegota S, Lorencin I, Ohkura K, Car Z. On the Traveling Salesman Problem in Nautical Environments: an Evolutionary Computing Approach to Optimization of Tourist Route Paths in Medulin, Croatia. *Pomorski zbornik*. 2019 Dec 31;57(1):71-87.
10. Lorencin I, Anđelić N, Mrzljak V, Car Z. Marine Objects Recognition Using Convolutional Neural Networks. *NAŠE MORE: znanstveno-stručni časopis za more i pomorstvo*. 2019 Sep 26;66(3):112-9.
11. Berlyand Y, Raja AS, Dorner SC, Prabhakar AM, Sonis JD, Gottumukkala RV, et al. How artificial intelligence could transform emergency department operations. *The American journal of emergency medicine*. 2018 Aug 1;36(8):1515-7.
12. Cismondi F, Horn AL, Fialho AS, Vieira SM, Reti SR, Sousa JM, et al. Multi-stage modeling using fuzzy multi-criteria feature selection to improve survival prediction of ICU septic shock patients. *Expert Systems with Applications*. 2012 Nov 15;39(16):12332-9.
13. Crone SF, Lessmann S, Stahlbock R. The impact of preprocessing on data mining: An evaluation of classifier sensitivity in direct marketing. *European Journal of Operational Research*. 2006 Sep 16;173(3):781-800.
14. Chittineni S, Bhogapathi RB. A Study on the Behavior of a Neural Network for Grouping the Data. *arXiv preprint arXiv:1203.3838*. 2012 Mar 17.
15. Krenker A, Bešter J, Kos A. Introduction to the artificial neural networks. *Artificial Neural Networks: Methodological Advances and Biomedical Applications*. InTech. 2011 Apr:1-8.
16. Mohamed AR, Dahl GE, Hinton G. Acoustic modeling using deep belief networks. *IEEE transactions on audio, speech, and language processing*. 2011 Jan 31;20(1):14-22.
17. Clevert DA, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (elus). *arXiv preprint arXiv:1511.07289*. 2015 Nov 23.
18. Suykens JA, Vandewalle J. Least squares support vector machine classifiers. *Neural processing letters*. 1999 Jun 1;9(3):293-300.
19. Nalbantov GI, Bioch JC, Groenen PJ. Classification with support hyperplanes. In *European Conference on Machine Learning 2006 Sep 18 (pp. 703-710)*. Springer, Berlin, Heidelberg.
20. Mukherjee S, Osuna E, Girosi F. Nonlinear prediction of chaotic time series using support vector machines. In *Neural Networks for Signal Processing VII. Proceedings of the 1997 IEEE Signal Processing Society Workshop 1997 Sep 24 (pp. 511-520)*. IEEE.
21. Yu H, Jiang X, Vaidya J. Privacy-preserving SVM using nonlinear kernels on horizontally partitioned data. In *Proceedings of the 2006 ACM symposium on Applied computing 2006 Apr 23 (pp. 603-610)*.
22. Denoeux T. A k-nearest neighbor classification rule based on Dempster-Shafer theory. In *Classic works of the Dempster-Shafer theory of belief functions 2008 (pp. 737-760)*. Springer, Berlin, Heidelberg.
23. Zheng Z, Su D. Short-term traffic volume forecasting: A k-nearest neighbor approach enhanced by constrained linearly sewing principle component algorithm. *Transportation Research Part C: Emerging Technologies*. 2014 Jun 1;43:143-57.
24. Tao Y, Papadias D, Lian X. Reverse kNN search in arbitrary dimensionality. In *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30 2004 Aug 31 (pp. 744-755)*. VLDB Endowment.
25. Wang XH, Liu A, Zhang SQ. New facial expression recognition based on FSVM and KNN. *Optik*. 2015 Nov 1;126(21):3132-4.
26. Halligan S, Altman DG, Mallett S. Disadvantages of using the area under the receiver operating characteristic curve to assess imaging tests: a discussion and proposal for an alternative approach. *European radiology*. 2015 Apr 1;25(4):932-9.
27. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. *Indian pediatrics*. 2011 Apr 1;48(4):277-87.
28. Narendra KS, Mukhopadhyay S. Neural networks in control systems. In [1992] *Proceedings of the 31st IEEE Conference on Decision and Control 1992 Dec 16 (pp. 1-6)*. IEEE.
29. Ghosh S, Li J, Cao L, Ramamohanarao K. Septic shock prediction for ICU patients via coupled HMM walking on sequential contrast patterns. *Journal of biomedical informatics*. 2017 Feb 1;66:19-31.

30. Guo G, Wang H, Bell D, Bi Y, Greer K. KNN model-based approach in classification. In "OTM Confederated International Conferences" On the Move to Meaningful Internet Systems" 2003 Nov 3 (pp. 986-996). Springer, Berlin, Heidelberg.
31. Wang SL, Wu F, Wang BH. Prediction of severe sepsis using SVM model. In "Advances in computational biology 2010" (pp. 75-81). Springer, New York, NY.
32. NirmalaDevi M, alias Balamurugan SA, Swathi UV. An amalgam KNN to predict diabetes mellitus. In "2013 IEEE International Conference ON Emerging Trends in Computing, Communication and Nanotechnology (ICECCN) 2013 Mar 25" (pp. 691-695). IEEE.
33. Brnić M, Čondrić E, Blažević S, Anđelić N, Borović E, Car Z. Sepsis prediction using artificial intelligence algorithms. In "International Conference on Innovative Technologies, IN-TECH 2018" 2018 Jan 1.