# Pancreatic cancer detection with self-expanding ART fuzzy neural network online training

**André Luiz Caliari Costa[1], Reginaldo José da Silva[2], Mara Lúcia Martins Lopes[3] and Angela Leite Moreno[4]**

## ABSTRACT

Pancreatic cancer is one of the most aggressive cancers and has a high mortality rate due to late diagnosis, as the first symptoms usually appear when it is at an advanced stage, making resection of the tumor impossible. With the aim of assisting in early diagnosis, this article presents the results obtained using urinary biomarkers to detect pancreatic cancer. The Self-Expanding ART Fuzzy Neural Network Online Training achieved superior results compared to other machine learning models in the different types of proposed classifications. This makes the use of our method promising to improve the diagnosis of pancreatic cancer compared to the methods still used and contributing to precision medicine.

**Keywords:** Pancreatic Ductal Adenocarcinoma, Adaptive Resonance Theory, Machine Learning, Fuzzy ART.

[1] Master's degree, USP – University of São Paulo, Ribeirão Preto, São Paulo, Brazil.
[2] Master's degree, UNESP – São Paulo State University "Júlio de Mesquita Filho", Ilha Solteira, São Paulo, Brazil.
[3] Postdoc, UNESP – São Paulo State University "Júlio de Mesquita Filho", Ilha Solteira, São Paulo, Brazil.
[4] Doctorate, UNIFAL – University Federal of Alfenas, Alfenas, Minas Gerais, Brazil.

## INTRODUCTION

Pancreatic Ductal Adenocarcinoma (PDAC) is the most common pancreatic cancer, accounting for 90% of pancreatic tumors. It is also the most aggressive, with high morbidity and mortality rates. The 5-year survival rate for is approximately 8%, which is a relatively low rate compared to other types of tumors (Ardengh, Coelho, Osvaldt, 2008; Ferlay et al. 2015).

The main risk factors for ADP involve environmental and hereditary factors. Environmental factors include smoking, obesity, and diabetes mellitus, while hereditary factors include hereditary breast and ovarian cancer associated with the BRCA1 and BRCA2 genes, Peutz-Jeghers syndrome, and hereditary pancreatitis. Age is also a risk factor, as 80% of patients with Pancreatic Ductal Adenocarcinoma are diagnosed between the ages of 60 and 80 (Ardengh et al., 2008).

The most common symptoms of PAD are weakness, weight loss, lack of appetite, abdominal pain, jaundice, nausea and back pain. Due to the lack of specific symptoms, the delay in presenting the first symptoms makes early diagnosis impossible (Ardengh et al., 2008).

When cancer is diagnosed in the early stages, it is possible to start treatment to prevent the cancer from progressing. The main form of cancer treatment is surgery, without which there is no prospect of a cure. However, diagnosis tends to occur when the disease is at an advanced stage, making it impossible to resect the tumor (Cooperman, Iskandar, Wayne & Steele , 2018).

Diagnosis is typically based on common symptoms, and imaging tests like computed tomography (CT) and ultrasound, with CT being the primary diagnostic tool (Cooperman et al., 2018).

Certain genes associated with cancer show alterations in their expression levels, either being underexpressed or overexpressed, compared to non-cancer patients. These genes can serve as biomarkers for diagnosing cancer and planning appropriate therapies. There are thousands of protein-coding genes that have been identified as potential candidates for pancreatic cancer biomarkers (Harsha et al. 2009). Discovering genes specific to pancreatic cancer or creating a panel of biomarkersfor early diagnosis poses a significant challenge (Lennon et al., 2014; Capello et al., 2017; Singhi et al., 2019). Based on previous studies, we conducted simulations focusing on three protein biomarkers excreted in human urine. These biomarkers can differentiate between pancreatic cancer patients and healthy individuals. (Radon et al., 2015).

This study aims to present the results obtained using a neural network based on Adaptive Resonance Theory (ART) for online training in diagnosing patients with ADP using urinary biomarkers, creatinine, and age. The data used is available in Debernardi et al. (2020). Additionally, it includes a comparison with results from the literature and some of the most commonly used classification models.

## ADAPTIVE RESONANCE THEORY

This theory, proposed by Grossberg (1976), is based on the cognitive theory of how the brain categorizes, recognizes, and predicts objects. It aims to address the balance between stability and plasticity in neural networks. Plasticity allows the network to adapt and group input patterns, while stability enables the network to learn new patterns without forgetting previously acquired knowledge (Carpenter & Grossberg, 1987a).

Several networks based on this theory have been developed. The pioneering architecture using this theory was the Grossberg Network (GN) designed by Grossberg in 1976. Since then, several other networks, such as ART1 have been created (Carpenter & Grossberg, 1987a), ART2 (Carpenter & Grossberg, 1987b), ART Fuzzy (Carpenter, Grossberg & Rosen, 1991c), ARTMAP (Carpenter, Grossberg & Reynolds., 1991b) and ARTMAP Fuzzy (Carpenter, Grossberg, Markuzon & Reynolds, 1992).

In general, ART networks consist of two main subsystems (Carpenter & Grossberg, 1991a). The first subsystem is the attention subsystem, which comprises three layers of interconnected neurons. The $F_0$ layer is responsible for preprocessing, where input data is normalized and complemented to ensure uniform magnitude across all input vectors (Moreno, 2010). The $F_1$ layer receives data from the $F_0$ layer and propagates it to the $F_2$ units. The $F_2$ layer groups training patterns into recognition categories. The second subsystem, the guidance subsystem, resonates and resets using the surveillance parameter ($\rho$) to determine if an input pattern can be included in existing categories (Moreno, 2010). The basic architecture of an ART network is illustrated in Figure 1.



Figure 1. ART architecture.
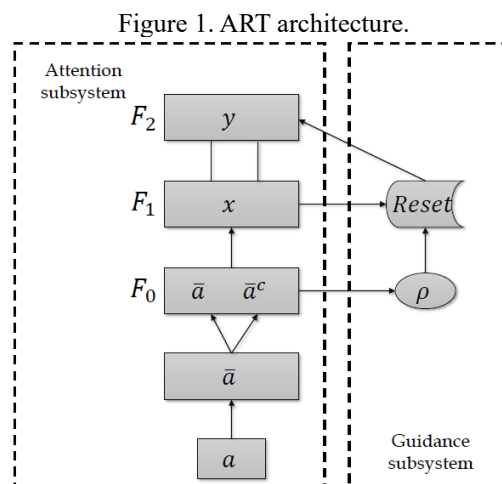
All these fields are interconnected through weights, which store information by selecting categories, equalization criteria, and training (Carpenter et al., 1991c). The connections between layers F1 and F2 are non-recurrent or feedforward connections, where each neuron in F1 is connected to all neurons in F2. Conversely, layers F2 to F1 are linked by recurrent connections

(feedback), with each neuron in F2 connected to all neurons in F1 (Carpenter et al., 1991c). The key parameters in an ART neural network include:

- α: Learning rate (α > 0), the parameter responsible for controlling the influence that new patterns on the cluster.
- β: Training rate (0 < β ≤ 1) determines the speed of training. A value of β = 1 indicates fast training, while β < 1 indicates slow training.
- ρ: Surveillance parameter (0 < ρ ≤ 1) is a control parameter that determines the level of similarity required for a new cluster to be created.

## SELF-EXPANDING FUZZY ART ONLINE TRAINING

Developed by Moreno (2016), the Self-Expanding ART Fuzzy Neural Network Online Training (SEARTFOT) is built upon Carpenter and Grossberg's ART Fuzzy model from 1991. These networks utilize fuzzy set theory by Zadeh (1965), using the AND operator (Λ) as defined by:

$$(x \wedge y) = min\{x_i, y_i\} \qquad\qquad (1)$$

The network proposed by Moreno (2016) is self-expanding and does not have a fixed structure. It initiates the training process by selecting the first pattern as the center of the first cluster and expands as needed. This modification enhances the network's accuracy, reduces processing time, and lowers computational costs by only comparing the similarity of active clusters. Additionally, the network features an online training module that enables continuous learning, improving efficiency in prediction and training. The integration of this module in ART networks is feasible due to their stability and plasticity. With online learning, there is no requirement to restart training when new data is introduced, allowing new patterns to be permanently added to the network's memory.

This network has temporary and permanent categories to prevent rare data from being stored in the network's memory. Two new parameters have been introduced: the permanence parameter (NMIN) determines when a temporary category becomes permanent, and the novelty index (η) controls the updating of weights for permanent categories, filtering out irrelevant information (Moreno, 2010) to prevent memory overload. Resonance occurs when a pattern belongs to a temporary cluster, while resonance in a permanent category only happens if the input is close enough to the category's center within the parameter η. The Figure 2 shows the pseudocode for the ART Fuzzy Self-Expanding Network Online Training (SEARTFOT).

Figure 2: Pseudocode of the Self-Expanding ART Fuzzy Network Online Training.

**Input:** $a$: Input Data;
$\alpha$: Learning Rate $(\alpha > 0)$;
$\beta$: Training Parameter $(0 < \beta \leq 1)$;
$\rho$: Vigilance Parameter $(0 < \rho \leq 1)$;
$\eta$: Novelty Index;
$NMIN$: Parameter of permanence.
**Output:** $y^{F_2}$: Activity vector

1: **procedure**

2: $\quad Cluster_{temp} = 1, \quad \overline{a} = \dfrac{a}{\left|\sum\limits_{i}^{M} a_i\right|}, \quad \overline{a}_i^c = 1 - \overline{a}_i, \quad I = [\overline{a} \quad \overline{a}^c]$

3: $\quad W_1^T = I_1 \quad ConT_1 = 1$

4: $\quad$ **while** Data for training exists **do**

5: $\quad\quad Counter = 1$

6: $\quad\quad Cluster_{temp} + Cluster_{def} = Categories$

7: $\quad\quad$ **for** all temporary and permanent clusters **do**

8: $\quad\quad\quad T_j^T = \dfrac{|I \wedge W_j^T|}{\alpha + |W_j^T|}, \quad\quad T_j^D = \dfrac{|I \wedge W_j^D|}{\alpha + |W_j^D|}$

9: $\quad\quad flag = True$

10: $\quad\quad$ **while** flag = True **do**

11: $\quad\quad\quad T_J = \max\{T_j^T, T_j^D\}$

12: $\quad\quad\quad M_J = \dfrac{|I \wedge W_J^T|}{|I|} \quad or \quad M_j = \dfrac{|I \wedge W_J^D|}{|I|}$

13: $\quad\quad\quad$ **if** $M_J \geq \rho$ **then**

14: $\quad\quad\quad\quad flag = False$

15: $\quad\quad\quad\quad$ **if** $T_J^D > T_J^T$ **then**

16: $\quad\quad\quad\quad\quad$ **if** $T_J^D > \eta$ **then**

17: $\quad\quad\quad\quad\quad\quad W_J^D = \beta(I \wedge W_J^D) + (1-\beta)W_J^D$

18: $\quad\quad\quad\quad$ **else**

19: $\quad\quad\quad\quad\quad W_J^T = \beta(I \wedge W_J^T) + (1-\beta)W_J^T$

20: $\quad\quad\quad\quad\quad ConT_J^{new} = ConT_J^{old} + 1$

21: $\quad\quad\quad\quad\quad$ **if** $ConT_J = NMIN$ **then**

22: $\quad\quad\quad\quad\quad\quad Cluster_{def} = Cluster_{def} + 1$

23: $\quad\quad\quad\quad\quad\quad W_{Cluster_{def}}^D = W_J^T$

24: $\quad\quad\quad\quad\quad\quad Cluster_{temp} = Cluster_{temp} - 1$

25: $\quad\quad\quad\quad\quad\quad W_J^T = W_{J+1}^T$

26: $\quad\quad\quad\quad\quad\quad ConT_J = ConT_{J+1}$

27: $\quad\quad\quad$ **else**

28: $\quad\quad\quad\quad$ **if** $Categories > Counter$ **then**

29: $\quad\quad\quad\quad\quad Counter^{new} = Counter^{old} + 1$

30: $\quad\quad\quad\quad\quad T_J = 0$

31: $\quad\quad\quad\quad$ **else**

32: $\quad\quad\quad\quad\quad Cluster_{temp}^{new} = Cluster_{temp}^{old} + 1$

33: $\quad\quad\quad\quad\quad J = Cluster_{temp}$

34: $\quad\quad\quad\quad\quad W_J^T = I_i$

35: $\quad\quad\quad\quad\quad flag = False$

36: $\quad\quad y_i = \begin{cases} 1, & i \neq J, \\ 0, & i = J. \end{cases}$

## METHODOLOGY

### MACHINE LEARNING MODELS

In this study, we compared the results of the ARTFAEOT network with seven commonly used machine learning methods. These methods, implemented in Python using the Scikit-Learn library, included Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting (GB), MultiLayer Perceptron (MLP), and XGBoost (XGB). The optimal hyperparameters for each model were determined using the GridSearchCV method, which exhaustively tests all possible combinations of hyperparameters provided in a grid search.

### DATASET

The data used in this article was collected and made available by Debernardi et al. (2020). It consists of 590 samples, including 183 from healthy individuals without pancreatic diseases, malignancies, or kidney disease history, 208 from patients with benign hepatobiliary diseases, and 199 from patients with autoimmune pancreatitis (ADP) at various stages (I-IV). The benign group comprised 119 cases of chronic pancreatitis, 54 samples from patients with gallbladder diseases, 20 with cystic lesions of the pancreas, and 15 cases with abdominal pain and gastrointestinal symptoms suggestive of pancreatic origin. The main attributes and sample quantities are shown in Table 1.

Table 1. Main attributes of the data set.

| Attributes | Number of samples |
|---|---|
| Age | 590 |
| Diagnosis | 590 |
| Plasma CA19-9 U/ml | 350 |
| Creatinine mg/ml | 590 |
| LYVE1 ng/ml | 590 |
| REG1B ng/ml | 590 |
| TFF1 ng/ml | 590 |
| REG1A ng/ml | 306 |

The biomarkers Plasma CA19-9 and REG1A were excluded from the simulations due to missing values.

The three remaining protein biomarkers LYVE-1, REG1B, and TFF1 are commonly downregulated in men and women with pancreatic adenocarcinoma (Radon et al., 2015). The LYVE-1 (Lymphatic Vessel Endothelial Receptor 1) protein is a transmembrane receptor for the extracellular matrix glycosaminoglycan hyaluronan (HA) and is expressed on the luminal surface cells of the lymphatic endothelium, acting as a transporter of HA, mediating its uptake for catabolism within endothelial cells (Stanly et al., 2020). It is also associated with leukocyte trafficking, tumor progression, and metastasis, so the localization of HA targeted to LYVE-1 to lymphatic tissues may

provide a substrate for tumor cell migration, contributing to metastatic spread (Jackson, 2019; Jackson, 2022).

REG1B belongs to a family of proteins encoded by Reg genes (Moriizumi et al., 1994). Its up-regulation has been identified in precursor lesions of pancreatic adenocarcinoma, leading to accelerated cell proliferation and tumor growth (Li et al., 2016). This gene encodes the glycoprotein REG1B (regenerating family member 1 beta), which is secreted by the pancreas, primarily by pancreatic acinar cells, and is involved in the regeneration of pancreatic islet cells in patients with pancreatitis (Jin et al., 2011). REG1B is a more effective biomarker than its paralog REG1A in distinguishing between control samples and stage I-IIA pancreatic adenocarcinoma samples (Debernardi et al., 2020).

The TFF1 (Trefoil factor 1) gene belongs to the Trefoil Family Factor (TFF) genes and is primarily expressed in the mucosal cells of gastric surfaces. It encodes the TFF1 protein, previously known as pS2. TFF1 functions by protecting and enhancing the viscoelasticity of mucus in the gastrointestinal epithelium, where it is co-secreted with mucins. Its expression is upregulated during inflammatory conditions like gastrointestinal ulcers and pancreatitis, playing a key role in epithelial reconstitution. Moreover, TFF1 expression is observed in premalignant conditions in various tumors and serves as a prognostic marker in breast cancer and pancreatic adenocarcinoma.

The correlation coefficient was used to determine the relationship between the variables. Figure 3 displays the heat map of the Pearson correlation matrix for each attribute. Upon observation, it is evident that the attributes did not exhibit a significant correlation.

Figure 3: Pearson's correlation coefficient.
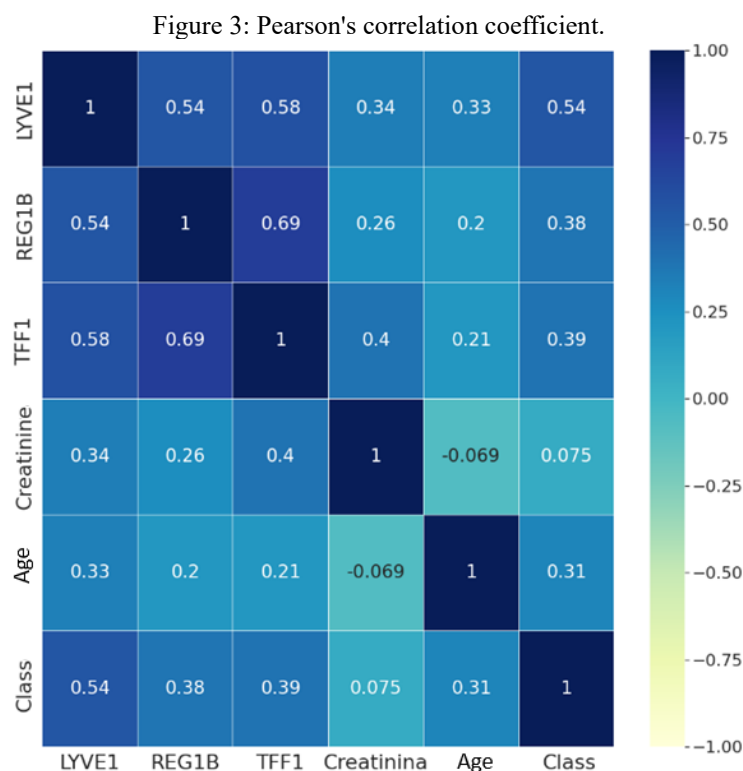
## EVALUATION METRICS

The confusion matrix, also known as the error matrix, was utilized to calculate the evaluation metrics in this article, including accuracy, sensitivity, and specificity. This matrix helps visualize the classification results by showing the number of correct classifications compared to the predicted ones. It also illustrates the relationship between two types of errors: type I error (false positive) and type II error (false negative).

Table 2. Confusion Matrix.

|          | True | False |
|----------|------|-------|
| Positive | TP   | FN    |
| Negative | FP   | TN    |

on what:

- TP: positive data correctly classified;
- TN: negative data classified correctly;
- FP: negative data classified incorrectly;
- FN: positive data classified incorrectly.

The following metrics are used:

- Accuracy (ACC): assesses the overall performance of the model, which is the most common measure for quantifying the performance of models.

$$ACC = \frac{VP + VN}{VP + VN + FP + FN} \tag{2}$$

- Sensitivity (Se): indicates the model's ability to detect truly positive individuals, i.e. the probability of a test being positive in those who are actually ill.

$$Se = \frac{VP}{VP + FN} \tag{3}$$

- Specificity (Sp): this is a complementary measure to sensitivity and indicates the model's ability to detect true negatives, i.e. it indicates the probability of a test being negative in those who are not actually ill.

$$Sp = \frac{VN}{VN + FP} \tag{4}$$

The data was partitioned using the Hold-Out repeated cross-validation method, also known as Monte Carlo cross-validation. This method involves randomly dividing the data set into two parts multiple times. Each division uses a 50-50 proportion, with one part as the training set and the other as the prediction set. The cross-validation process was repeated 50 times for all models, ensuring random shuffling and division of the data. (Dudoit & van der Laan, 2005; Borra & Di Ciaccio, 2010).

The Self-Expanding ART Fuzzy Neural Network Online Training was implemented using the Scilab computer system, version 6.0.1.

## RESULTS AND DISCUSSION

The simulations were divided into three stages, following the methodology of Debernardi et al. (2020). In the first stage, the classification was performed between healthy patients (control group) and patients with ADP, with a total of 382 samples. The accuracy, sensitivity, and specificity achieved by the SEARTFOT network in the 50 iterations of the Monte Carlo cross-validation method for this stage are shown in Figure 4.

Figure 4 - Accuracy obtained by the SEARTFOT network in classifying the control and ADP groups.
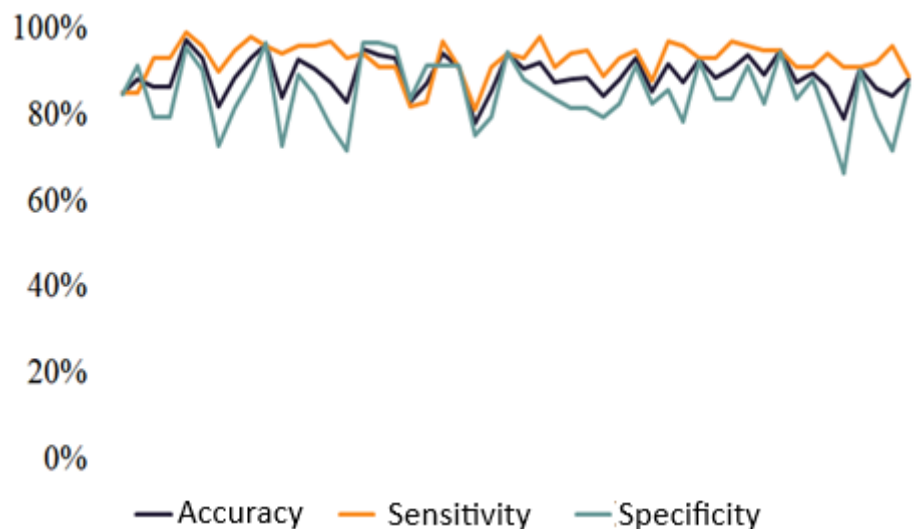


The average performance of each model is shown in Table 3. For this problem, high sensitivity is crucial to accurately diagnose cancer patients and avoid type II errors, where a cancer patient is misdiagnosed as healthy. While the SVM model had high sensitivity, its low specificity (less than 20%) indicates poor performance as it labeled most samples as cancer patients. This is evident in the Area Under the Curve (AUC) analysis, with the model achieving $0.582 \pm 0.015$, showing an inability to distinguish between classes.

Table 3. Results obtained by each of the models tested.

| Method | ACC (%) ± Deviation | Se (%) ± Deviation | Sp (%) ± Deviation |
|---|---|---|---|
| ARTFAEOT | 88.95 ± 4.36 | 92.68 ± 4.11 | 84.89 ± 7.41 |
| RL | 84.26 ± 1.95 | 80.83 ± 3.47 | 88.13 ± 3.54 |
| KNN | 79.51 ± 2.40 | 76.65 ± 3.30 | 82.62 ± 5.26 |
| XGB | 85.66 ± 2.00 | 87.30 ± 3.00 | 83.89 ± 4.08 |
| RF | 84.57 ± 1.72 | 84.34 ± 4.14 | 84.81 ± 3.67 |
| GB | 84.90 ± 2.20 | 86.64 ± 3.11 | 83.01 ± 4.39 |
| SVM | 59.81 ± 1.62 | 97.87 ± 0.94 | 18.41 ± 3.46 |
| MLP | 79.37 ± 2.93 | 79.05 ± 3.85 | 79.72 ± 5.12 |

The logistic regression model showed good specificity but had a lower average AUC of 0.843 ± 0.018 compared to the SEARTFOT network, which had an average AUC of 0.892 ± 0.045. The SEARTFOT network demonstrated superior performance in distinguishing between healthy and ADP patients, surpassing all other models evaluated. Figure 5 displays the AUC values for each of the 50 iterations in the Monte Carlo validation method.

Figure 5: AUC obtained by the SEARTFOT network and RL in the classification between the control group and ADP.
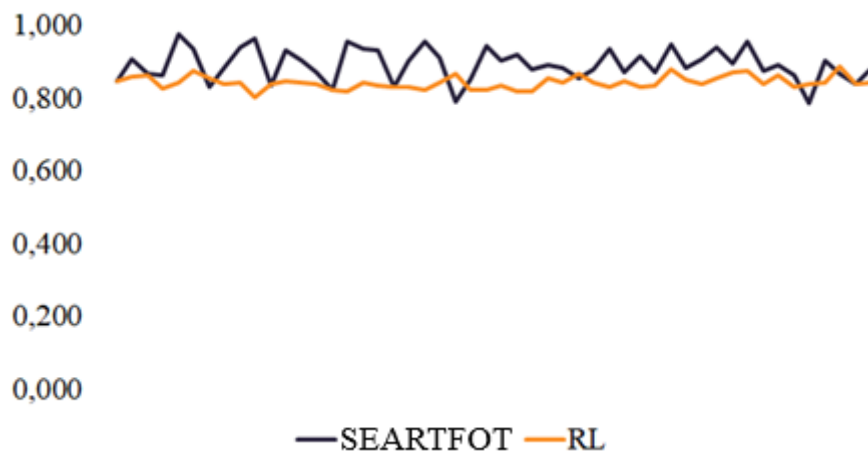


The second stage involved classifying patients with benign hepatobiliary diseases (benign group) and patients with ADP, totaling 407 samples. The accuracy, sensitivity, and specificity achieved by the SEARTFOT network in the 50 iterations of the Monte Carlo cross-validation method for this stage are shown in Figure 6.

Figure 6. Accuracy obtained by the SEARTFOT network in classifying the benign and ADP groups.
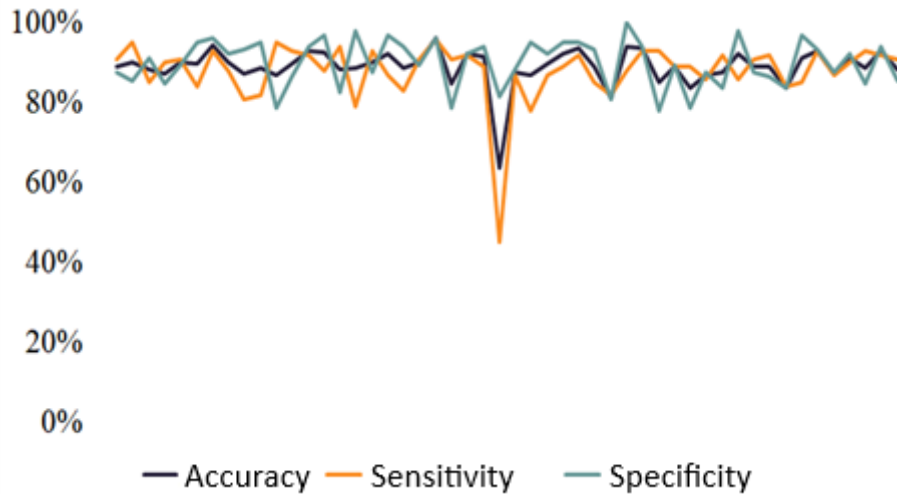


The accuracy, sensitivity, and specificity results for classifying patients with benign hepatobiliary diseases and ADP using the methods in this study are presented in Table 4. The SEARTFOT network consistently outperformed other methods, particularly in terms of AUC, achieving a value of $0.891 \pm 0.047$. This outperformed the AUC of 0.849 obtained by the PancRISK (Logistic Regression) method used by Debernardi et al. (2020). Figure 7 illustrates the AUC values for the SEARTFOT and RL models across 50 iterations of the Monte Carlo cross-validation method.

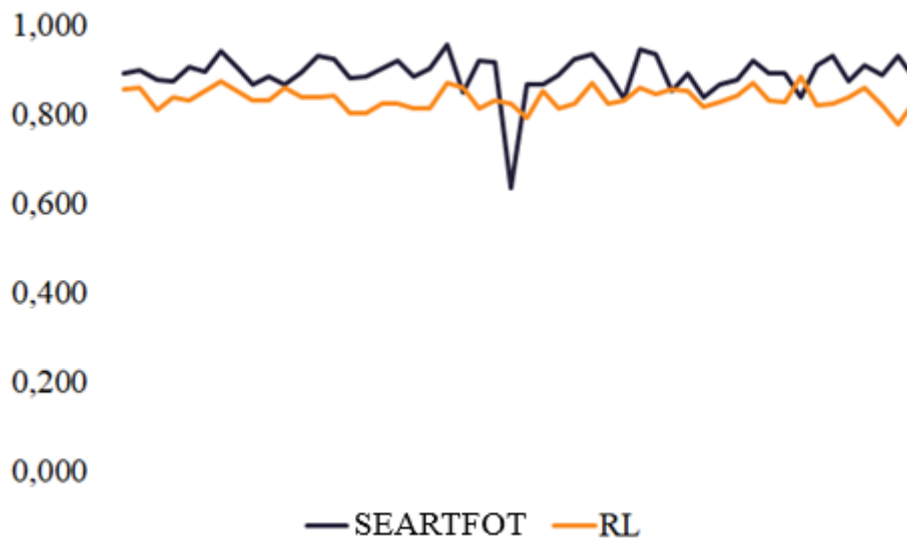Figure 7: AUC obtained by the SEARTFOT network and RL in the classification between benign and ADP groups.

Table 4: Results obtained from the models tested.

| Method | ACC (%) $\pm$ deviation | Se (%) $\pm$ deviation | Sp (%) $\pm$ deviation |
|---|---|---|---|
| SEARTFOT | 89.06 $\pm$ 4.72 | 88.04 $\pm$ 7.53 | 90.04 $\pm$ 5.89 |
| RL | 75.40 $\pm$ 2.09 | 72.34 $\pm$ 5.13 | 78.35 $\pm$ 4.23 |
| KNN | 69.24 $\pm$ 2.67 | 67.30 $\pm$ 5.37 | 71.10 $\pm$ 5.12 |
| XGB | 75.41 $\pm$ 2.56 | 74.82 $\pm$ 5.60 | 75.98 $\pm$ 4.70 |
| RF | 75.20 $\pm$ 2.19 | 74.24 $\pm$ 4.91 | 76.12 $\pm$ 5.19 |
| GB | 73.13 $\pm$ 2.47 | 72.36 $\pm$ 5.37 | 73.87 $\pm$ 3.94 |
| SVM | 65.28 $\pm$ 2.39 | 54.16 $\pm$ 7.59 | 75.96 $\pm$ 6.74 |
| MLP | 73.45 $\pm$ 3.98 | 69.50 $\pm$ 6.19 | 77.25 $\pm$ 5.12 |

Table 5 compares the results of the SEARTFOT network with the best results from Debernardi et al. (2020). The network in this study achieved higher sensitivity compared to the PancRISK (Logistic Regression) model. It is important to note that a test with high sensitivity is preferred.

Table 5. Comparison of classification results between patients in the control group and ADP.

| | Method | Se (%) | Sp (%) |
|---|---|---|---|
| This work | SEARTFOT | 92.68 | 84.89 |
| Debernardi et al. (2020) | PancRISK (Logistic Regression) | 85.00 | 82.80 |

When comparing the results obtained by the SEARTFOT network with PancRISK (Logistic Regression) used by Debernardi et al. (2020) for classifying patients with benign hepatobiliary diseases and ADP, the network in this study demonstrated superior average performance. It achieved a sensitivity of 88.04% and specificity of 90.04%, as shown in Table 6.

Table 6. Comparison of classification results between patients with benign hepatobiliary diseases and ADP.

| | Method | Se (%) | Sp (%) |
|---|---|---|---|
| This work | SEARTFOT | 88.04 | 90.04 |
| Debernardi et al. (2020) | PancRISK (Logistic Regression) | 72.50 | 80.00 |

The third stage involved classifying patients with chronic pancreatitis (CP) and autoimmune pancreatitis (ADP), using 318 samples (119 CP and 199 ADP). The network's performance exceeded that of Debernardi et al. (2020), as shown in Table 7.

Table 7. Comparison of classification results between patients with CP and ADP.

| | Method | Se (%) | Sp (%) |
|---|---|---|---|
| This work | SEARTFOT | 84.20 | 86.50 |
| Debernardi et al. (2020) | PancRISK (Logistic Regression) | 80.00 | 67.60 |

## CONCLUSION

In this study, various machine-learning models were used to identify patients with pancreatic cancer. The ART Fuzzy Self-Expanding Online Training network outperformed other models, achieving an average accuracy of 88.95% ± 4.36% in distinguishing between cancer patients and

healthy individuals. When differentiating between cancer patients and those with benign hepatobiliary diseases, the network achieved an average accuracy of $89.06\% \pm 4.72\%$, demonstrating its effectiveness in pancreatic cancer detection.

The ART Fuzzy Self-Expanding Online Training network and biomarkers showed promising results for this issue, indicating their potential to aid medical professionals in decision-making and enabling efficient and early diagnosis.

## ACKNOWLEDGEMENTS

# REFERENCES

1. Ardengh, J. C., Coelho, N., & Osvaldt, A. B. (2008). Early pancreatic cancer: the methods currently available can identify this disease?. *Arquivos de Gastroenterologia, 45*, 169-177. (In Portuguese).

2. Borra, S., & Di Ciaccio, A. (2010). Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. *Computational statistics & data analysis, 54*(12), 2976-2989.

3. Capello, M., et al. (2017). Sequential validation of blood-based protein biomarker candidates for early-stage pancreatic cancer. *JNCI: Journal of the National Cancer Institute, 109*(4).

4. Carpenter, G. A., Grossberg, S., Markuzon, N., & Reynolds, J. H. (1992). Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on neural networks, 3*(5), 698-713.

5. Carpenter, G. A., & Grossberg, S. (1987a). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer vision, graphics, and image processing, 37*(1), 54-115.

6. Carpenter, G. A., & Grossberg, S. (1987b). ART 2: Self-organization of stable category recognition codes for analog input patterns. *Applied optics, 26*(23), 4919-4930.

7. Carpenter, G. A., & Grossberg, S. (1991a). Pattern Recognition by Self-organizing Neural Networks. *MIT Press, Cambridge*.

8. Carpenter, G. A., Grossberg, S., & Reynolds, J. H. (1991b). ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural networks, 4*(5), 565-588.

9. Carpenter, G. A., Grossberg, S., & Rosen, D. B. (1991c). Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural networks, 4*(6), 759-771.

10. Cooperman, A. M., Iskandar, M. E., Wayne, M. G., & Steele, J. G. (2018). Prevention and early detection of pancreatic cancer. *Surgical Clinics, 98*(1), 1-12.

11. Debernardi, S., et al. (2020). A combination of urinary biomarker panel and PancRISK score for earlier detection of pancreatic cancer: A case–control study. *PLoS Medicine, 17*(12), e1003489.

12. Dudoit, S., & van der Laan, M. J. (2005). Asymptotics of cross-validated risk estimation in estimator selection and performance assessment. *Statistical methodology, 2*(2), 131-154.

13. Emidio, N. B., et al. (2019). Trefoil factor family: Unresolved questions and clinical perspectives. *Trends in biochemical sciences, 44*(5), 387-390.

14. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D., & Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer, 136*(5), E359-E386.

15. Grossberg, S. (1976). Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological cybernetics, 23*(3), 121-134.

16. Harsha, H. C., et al. (2009). A compendium of potential biomarkers of pancreatic cancer. *PLoS medicine, 6*(4), e1000046.

17. Jackson, D. G. (2019). Hyaluronan in the lymphatics: the key role of the hyaluronan receptor LYVE-1 in leucocyte trafficking. *Matrix Biology, 78*, 219-235.

18. Jackson, D. G. (2022). Immune Cell Trafficking in the Lymphatics, Hyaluronan Biology and Tumour Metastasis. In: Cancer Metastasis Through the Lymphovascular System. *Springer, Cham*, 231-240.

19. Jin, C. X., et al. (2011). Pancreatic stone protein/regenerating protein family in pancreatic and gastrointestinal diseases. *Internal Medicine, 50*(15), 1507-1516.

20. Lennon, A. M., et al. (2014). The early detection of pancreatic cancer: what will it take to diagnose and treat curable pancreatic neoplasia?. *Cancer research, 74*(13), 3381-3389.

21. Li, Q., et al. (2016). Reg proteins promote acinar-to-ductal metaplasia and act as novel diagnostic and prognostic markers in pancreatic ductal adenocarcinoma. *Oncotarget, 7*(47), 77838-77853.

22. Moreno, A. L. (2010). Transient stability analysis using a modified Euclidean fuzzy Art-Artmap neural network with continuous training. PhD thesis, *São Paulo State University, Ilha Solteira*. (In Portuguese)

23. Moreno, A. L. (2016). ART and ARTMAP Neural Networks with Continuous Training: A detailed presentation of the main types of ART and ARTMAP neural networks. *Novas Edições Acadêmicas, Saarbrücken*. (In Portuguese).

24. Moriizumi, S., et al. (1994). Isolation, structural determination and expression of a novel reg gene, human reg Iβ. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression, 1217*(2), 199-202.

25. Radon, T. P., et al. (2015). Identification of a three-biomarker panel in urine for early detection of pancreatic adenocarcinoma. *Clinical Cancer Research, 21*(15), 3512-3521.

26. Singhi, A. D., et al. (2019). Early detection of pancreatic cancer: opportunities and challenges. *Gastroenterology, 156*(7), 2024-2040.

27. Stanly, T. A., et al. (2020). The cortical actin network regulates avidity-dependent binding of hyaluronan by the lymphatic vessel endothelial receptor LYVE-1. *Journal of Biological Chemistry, 295*(15), 5036-5050.

28. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*, 338-353.