Revista de Informática Teórica e Aplicada - RITA - ISSN 2175-2745
Vol. 31, Num. 2 (2024) 138-146

RESEARCH ARTICLE

Rough Sets and Multilayer Perceptron in Tandem for Processing the Aleatory Uncertainty in COVID-19 Cases

Rough Sets e Multilayer Perceptron em Conjunto para o Processamento da Incerteza para Casos de COVID-19

Kalyl Henings[1], Vinícius Hansen[1], Gilmário Barbosa dos Santos[2]

Abstract: It is common sense that challenging clinical cases can occur in the practice of medicine. These clinical cases can lead to undesirable situations of diagnostic uncertainty, making it important to identify these difficult cases and lead them to a discussion by experts for appropriate characterization. In remote/isolated regions, it is crucial to have a computational system that can support the health personnel in identifying these challenging cases. Events such as the COVID-19 pandemic have demonstrated the need for this type of system to assist in screening medical exams. Although the chest X-ray examination is a valuable diagnostic tool for COVID-19 cases, some conditions can be so challenging that doctors can be faced with uncertainty in diagnosis. This article proposes a system that combines machine learning via Multilayer Perceptron neural network in conjunction with the detection of uncertainty via Rough Sets in modeling a system that incorporates uncertainty to produce a classification of cases as positive for the disease, negative for the disease, or diagnostically uncertain. This system would serve as support for the rapid and efficient triage of cases, particularly those classified as “uncertain,” for a medical committee of specialists in a video conference, for example. Experiments were carried out and the trained model achieved 87.61% overall accuracy and a hit rate consistent across all classes.
Keywords: Unceartainty Treatment — Rough Sets — Multilayer Perceptron — COVID-19 — pulmonary RX

Resumo: É de conhecimento comum que na prática da medicina podem ocorrer casos clínicos desafiadores, os quais podem levar a situações indesejáveis de incerteza diagnóstica. É importante identificar tais casos e conduzi-los para uma discussão por especialistas buscando caracterizá-los apropriadamente. Especialmente em regiões mais remotas/isoladas, seria de grande valia a disponibilidade de um sistema computacional que pudesse apoiar os profissionais da área da saúde na identificação de casos desafiadores visando a devida triagem. Eventos como a pandemia de COVID-19 demonstraram a necessidade desse tipo de sistema para auxiliar na triagem de exames médicos. O exame de raio-X pulmonar tem sido valiosa ferramenta diagnóstica de casos de COVID-19, porém, em algumas condições podem ser muito desafiadoras expondo os médicos à incerteza sobre diagnósticos. Este artigo propõe um sistema que alia aprendizado de máquina via rede neural Multilayer Perceptron em conjunção com a detecção de incerteza via Rough Sets na modelagem de um sistema que incorpore a incerteza para produzir uma classificação de casos como positivo para a doença, negativo para a doença ou diagnóstico incerto. Esse sistema serviria como apoio para uma triagem rápida e eficiente de casos, particularmente os classificados como ”incertos”, para uma comissão médica de especialista em vídeo conferência, por exemplo. Experimentos foram realizados e o modelo treinado alcançou 87.61% de precisão geral e uma taxa de acerto consistente em todas as classes.

Palavras-Chave: Tratamento de Incerteza — Rough Sets — Multilayer Perceptron — COVID-19 — RX pulmonar

[1] *Computer Science Bachelor student, University of the State of Santa Catarina (UDESC), Rua Paulo Malschitzki, n. 200, Zona Industrial Norte, Joinville, Santa Catarina, CEP: 89.219-710, Brazil*
kalyl.henings,vinicius.hansen@edu.udesc.br
[2] *Computer Science Department, University of the State of Santa Catarina (UDESC), Rua Paulo Malschitzki, n. 200, Zona Industrial Norte, Joinville, Santa Catarina, CEP: 89.219-710, Brazil*
gilmario.santos@udesc.br
DOI: https://doi.org/10.22456/2175-2745.139289 • **Received:** 19/03/2024 • **Accepted:** 17/07/2024
CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introduction

The World Health Organization (WHO) has reported that the COVID-19 pandemic caused over 766 million confirmed cases and more than 6.9 million deaths globally [1]. These statistics underscore the pandemic's impact on public health systems, economies, and societal well-being around the world.

According to Baratella et al. [2], although computer tomography is the most sensitive imaging tool when coronavirus disease 2019 (COVID-19) is suspected, chest X-rays are frequently requested in patients with acute pulmonary symptoms. The authors [2] report the advantages of X-ray: it is inexpensive (particularly beneficial for poor countries), it is largely available at the bedside, and it produces low radiation exposure.

Given its wide availability throughout the world, the utilization of X-ray as a diagnostic tool is very important. X-ray imaging provides physicians with a valuable tool for screening internal structures, playing a critical role in the diagnosis of COVID-19 in humans [3].

Despite its significance, even experienced physicians may encounter instances of disagreement or uncertainty [4] on challenging cases exhibited on digital X-ray images.

The interpretation of X-ray images involves subjective analysis and relies on the expertise and experience of the radiologist or clinician. Complex anatomical variations, subtle abnormalities, overlapping structures, image artifacts, as well as doctors' level of expertise and fatigue, can contribute to the complexity of X-ray image interpretation, leading to differences in opinion among healthcare professionals. In such instances, second opinions and interdisciplinary consultations may enhance diagnostic accuracy.

To address this issue, this study proposes a model capable of identifying these challenging cases where uncertainty exists, thus serving as an auxiliary tool for medical professionals by performing a triage of clinical challenge cases in a way that saves resources and improves the logistics of examinations and diagnoses.

Considering the availability of a very large RX digital images dataset from COVID-19 screening containing cases diagnosed as positive or negative for COVID-19, since aleatoric uncertainty [5] [6] is inherent in data and usually affects doctors' diagnostics, the hypothesis is that this huge dataset will present uncertainty cases.

This paper proposes the use of Rough Sets (RS) to extract the uncertainty by mining this primary dataset, producing a new one with positive, negative, and uncertainty cases about COVID-19 diagnosis. This new dataset can be used for training and testing a Multilayer Perceptron Neural Network (MLP). This MLP will be capable of categorizing X-ray examinations into three distinct categories (positive, negative, and uncertainty cases about COVID-19 diagnosis). This scheme is depicted in Figure 1.

For clarification:

1. Uncertainty cases are difficult cases and cause doubt about diagnosis;
2. Supposing C_i and C_j are two similar cases in terms of respective X-ray image descriptors, they are part of the same "uncertainty" category if they present different diagnoses (2);
3. The necessary very large dataset was not available nor the computational power for processing it, so it is emphasized that this paper just reports a proof of concept shown in Figure 1.

Kurz et al. [5] maintain that there are two types of uncertainty: the one formally named *aleatoric uncertainty*, which is embedded in the data, and *epistemic uncertainty*, which is associated with the model that analyzed the data (usually a Deep Learning model).

In turn, Abdar and colleagues [6] assert that aleatoric uncertainty is an inherent property of the data distribution and, for that reason, it is irreducible.

It is important to highlight that this paper is focused on *aleatoric uncertainty*, and just for the sake of simplicity, the term *uncertainty* is used in this paper instead of *aleatoric uncertainty*.

The remainder of this paper is arranged as follows: in Section II, a discussion about related work is presented, followed by a general overview of the concept of the Multilayer Perceptron neural network in Section III. The foundations of the Rough Sets (RS) theory are presented in Section IV. In Section V, materials and methods of the experiments are given to illustrate the validity of the proposed pipeline, and results/discussion follow in Section VI.

2. Related Work

In a previous work about the use of open-source software in health systems [8], the authors reported the use of RS and a Support Vector Machine (SVM) for the triage of cases.

Despite the best efforts from the authors, based on the following five deep and important reviews, no other proposal similar to the one described in this paper was found. The reviews are listed as follows:

a) Kurz and colleagues[5] investigate uncertainty estimation applied to medical image classification. They also investigate metrics used to describe the effectiveness of the applied uncertainty estimation. They focused on the evaluation of model confidence, i.e., epistemic uncertainty;

b) Zoui et al. [7] review the various types of uncertainty, including aleatoric uncertainty and epistemic uncertainty, but they don't explore the processing of uncertainty using Rough Sets;

c) Huang and colleagues [9] weren't exclusively focused on probabilistic methods; this review also explores non-probabilistic approaches, such as Fuzzy Sets. Again, no exploration of Rough Sets was found in their review;

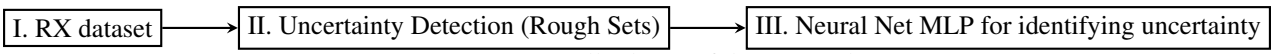


Figure 1. Overall scheme of the proposal.

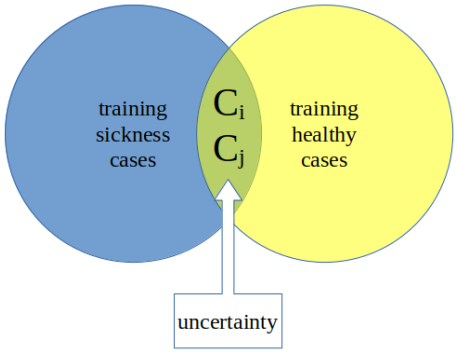


Figure 2. Aleatoric Uncertainty in data (adapted from [7]).

d) Alizadehsani and colleagues[10] conducted a review on a large collection of studies handling uncertainties in medical data using probability theory and machine learning techniques. So, their work was focused on aleatoric uncertainty. This review doesn't make any mention of Rough Sets;

e) Also the review of Abdar et al.[6] also discussed uncertainty quantification in general techniques, applications, and challenges. Although including Rough Sets, their approach doesn't include a neural network in tandem with Rough Sets.

3. Artificial Neural Networks & Multilayer Perceptron

Artificial neural networks (ANNs) are computational models inspired by the structure and functionality of biological neurons. ANNs originated from the pioneering work of Warren McCulloch and Walter Pitts in the 1940s, who aimed to create a mathematical model of the brain's neural networks. They introduced the concept of an artificial neuron, also known as a perceptron, which mimicked the behavior of biological neurons.

A perceptron can be compared to a neuron in terms of their basic functionality (Figure 3). Both the perceptron and the neuron receive inputs from multiple sources, apply weights to these inputs, and produce an output. However, there are notable differences. While a biological neuron has complex dendrites that receive inputs from thousands of other neurons, a perceptron typically receives a fixed number of inputs.

Biological neurons exhibit more diverse and intricate behaviors, such as temporal dynamics and synaptic plasticity, which are not typically captured by a simple perceptron model. Nevertheless, the concept of a perceptron captures the essence of a neuron's information processing capability and serves as a foundational unit in artificial neural networks.

A Multilayer Perceptron (MLP) is a type of feedforward artificial neural network that consists of one or more layers of nodes, or perceptrons, that process information and produce an output (Figure 4). The neurons in each layer are connected to the neurons in the next layer, forming a directed graph with information flowing from the input layer to the output layer. MLPs are commonly used in machine learning for supervised learning tasks such as classification and regression.

The training process involves adjusting the weights of the connections between the neurons in order to minimize a cost function that measures the difference between the predicted output and the desired output. MLPs are known for their ability to model complex nonlinear relationships between inputs and outputs and have been used in a wide range of applications, including image classification based on feature vectors.

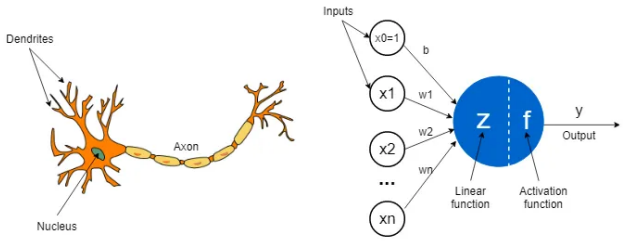


Figure 3. Neuron vs Perceptron

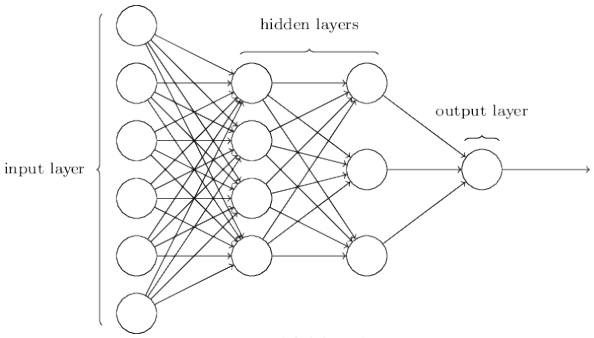


Figure 4. Two hidden layers MLP

3.1 Model Evaluation

The performance of the MLP neural network can be evaluated using various metrics such as Accuracy, Precision, Recall, F1 score, AUC (Area Under the Curve), and ROC (Receiver Operating Characteristic). A confusion matrix plot was generated to visualize the classification results

A confusion matrix is a summary table that shows the performance of a classification model by breaking down the predicted and actual class labels. It consists of True Positives, True Negatives, False Positives, and False Negatives. It helps evaluate the accuracy and error types of the model's

predictions. The AUC measures the model's ability to distinguish between positive and negative classes. A higher AUC indicates better performance, with 1 being perfect and 0.5 representing random guessing. It summarizes the model's performance across various thresholds.

ROC is a graphical plot that shows the performance of a classification model by plotting the true positive rate against the false positive rate at different threshold settings. It helps evaluate and compare models based on their ability to correctly classify positive and negative instances. The closer the curve is to the top-left corner, the better the model's performance. A higher AUC represents better discrimination power. All the metrics below can be extracted from the confusion matrix.

$$Accuracy = \frac{\text{Number of correct cases}}{\text{Total number of cases}} \tag{1}$$

$$Precision = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} \tag{2}$$

$$Recall = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}} \tag{3}$$

Micro and macro averaging approaches were taken for each metric. Macro-averaging calculates the performance for each class and takes the average across all classes. It computes the metric independently for each class, regardless of class imbalance, providing the performance of the classifier for each individual class. Micro-averaging calculates the performance metric considering the overall performance across all classes. It treats the classification task as a single, large set of predictions and calculates the metric based on the total number of true positives, false positives, and false negatives for all classes combined. It gives equal importance to each individual instance and emphasizes overall performance.

4. Rough Set Theory

Rough Sets is a mathematical framework that was introduced by Zdzislaw Pawlak [11] in the early 1980s as a way to deal with incomplete or uncertain information. It is based on the notion of approximation, where objects are classified based on their characteristics or attributes. Rough Sets provides a way to model and analyze data, and it has been applied in various fields, such as machine learning, data mining, decision-making, and pattern recognition. The theory provides concepts such as Upper-Approximation, Lower-Approximation, and Boundary.

The concept of Approximation in set theory refers to the classification of elements based on their level of inclusion in a given set. The Upper-Approximation represents elements that are completely included in the set A, while the Lower-Approximation represents elements that are partially included in both the set A and another set B. The Boundary of the set is derived as the difference between the Upper and Lower Approximations, which denotes those elements that have uncertain or ambiguous membership in both sets (Figure 5). Additional insights on this topic are available in the cited article [11]. Figure 5 was extracted from [12].

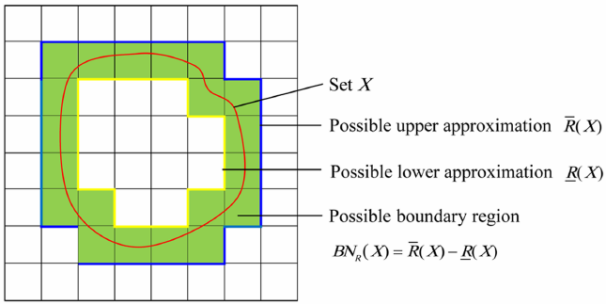


Figure 5. Rough Sets Concepts

According to Chen et al. [13], the classical Rough Sets model is mainly applicable to decision systems with discrete data, so it is not appropriate for real-valued datasets and is relevant to discrete-valued information systems [6].

5. Materials and Methods

This work was conducted on Fedora 37 in a Anaconda [14] environment with Python 3.10.11, on a Ryzen 7 3800X machine with 16GB of RAM. The dataset acquisition and preparation is described as follows:

5.1 Data Preparation

The use of the databases was compliant with ethical standards and regulations. All databases used anonymized data.

Two primary databases were processed in a sequence of procedures as shown in Figure 8 to get the "Final Dataset" used by the Neural Net (MLP). This serial processing is described as follows:

5.1.1 Primary Databases

Two primary databases were selected considering free access, documentation, and citations in the literature. They were the *COVID-19 Radiography* [15] [16] and the *BIMCV-COVID-19+* [17].

No image enhancement was performed, since the aim of the study is to generate a model capable of handling "raw" images.

5.1.2 Dataset-1

Initially, the two primary databases (*COVID-19 Radiography* and *BIMCV-COVID-19+*) were processed for the preparation of the so-called Dataset-1.

As seen in Table 1, the *COVID-19 Radiography* presents just 3,616 COVID-19 cases. On the other hand, it presents a significantly higher quantity of healthy cases, amounting to 10,200 healthy cases.

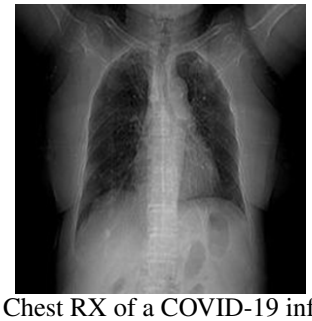


Figure 6. Chest RX of a COVID-19 infected lung

In order to complement the number of COVID-19 cases, a set of 6,576 cases was collected from the other primary database (*BIMCV*), amounting to 10,192 COVID-19 cases, as depicted in Figure 7.

Table 1. Primary Databases Selected

	Databases	
Cases	*COVID-19 Radiography*	*BIMCV*
COVID	3616	34829
Healthy	10200	-
Lung-Opacity	6012	-
Pneumonia	1345	-
Total	21173	34829

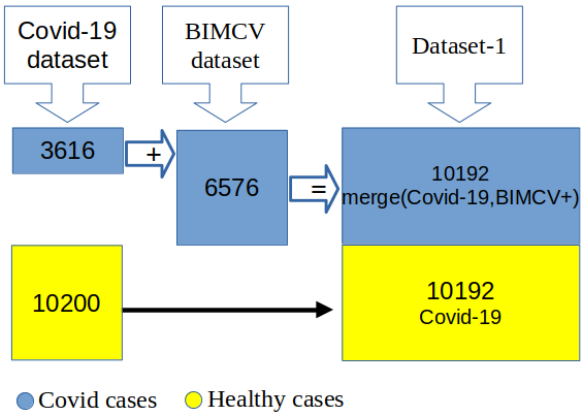


Figure 7. Assembling dataset-1 with equity in number of examples of cases. SSIM [18] [19] was applied for aiding the merging of *COVID-19 Radiography* and *BIMCV*

As seen in Table 2, Dataset-1 was composed of 20,384 cases distributed in 10,192 COVID cases and 10,192 healthy cases. The set of COVID cases results from the merge of cases from the two primary databases *COVID-19 Radiography* and *BIMCV* databases. All healthy cases in Dataset-1 were taken from the *COVID-19 Radiography* database; the *BIMCV* database does not contain cases of non-COVID patients.

Dataset-1 – Selecting Covid cases from *BIMCV*: In this section, the procedure using the Structural Similarity Index (SSIM) applied during the selection of 6,576 COVID cases from the *BIMCV* database is described.

Table 2. Dataset-1

Cases	*COVID-19 Radiography*	*BIMCV*	Total
Covid	3616	6576	10192
Healthy	10192	-	10192
Total	13808	6576	20384

SSIM was first introduced in 2004 [18] [19]. It is a method largely applied for image comparison. The algorithm consists of determining the similarity between two images simulating human perception, using the differences in luminance and contrast between the compared images. The description of how SSIM was applied here is explained as follows.

As described in Figure 7, 6,576 COVID cases were selected from *BIMCV* for merging with the COVID cases from the *COVID-19 Radiography* database. This selection needs to be cautious since the databases present images varying in terms of contrast, luminosity, artifacts, framing, and orientation.

The SSIM, via SciKit Image [20], was applied to select the 6,576 cases from *BIMCV* that presented similar profiles in comparison to cases in *COVID-19 Radiography*.

Dataset-1 went through the processes seen in Figure 8, which is described in detail in the subsequent sections.

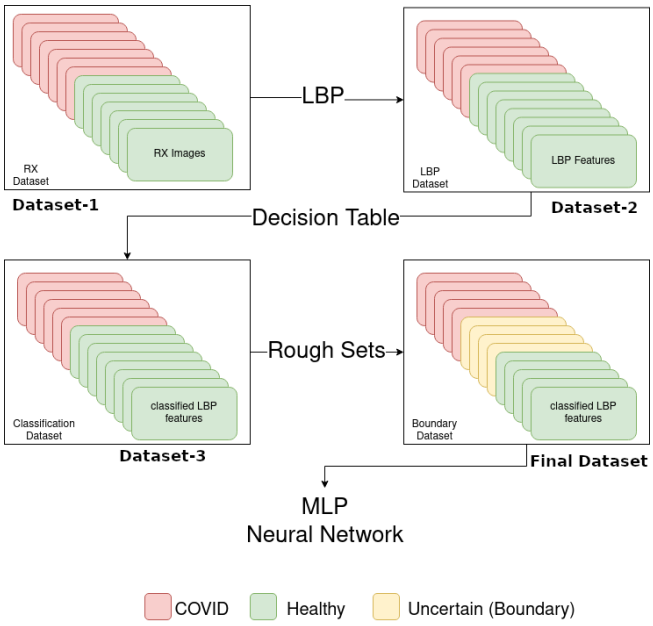


Figure 8. Data pipeline

5.1.3 Dataset-2 – Feature Extraction via LBP

The Local Binary Pattern (LBP) is a texture descriptor used in computer vision and image processing to describe texture properties. It encodes the relationship between a pixel and its neighboring pixels by comparing their gray-level values to a threshold [21]. The LBP descriptor has been widely used in

various applications, including in the context of COVID-19 chest X-rays [22].

The LBP method can be calibrated by the radius and number of points used to perform calculations of a pixel with its neighbors. The parameters used were Radius = 6 and Radius = 7, with the number of points = Radius * 8. According to this study [23], these are the best parameters for the context of chest X-ray. This method was applied to all images in the dataset before proceeding to the next stage.

Dataset-1 was processed by LBP, making Dataset-2 (as seen in Figure 8), which presents each case as an LBP descriptor of the respective RX.

5.1.4 Dataset-3 – Decision Table

The curation of the primary databases [15], [16], [17] was realized by prestigious medical institutions practicing the state of the art in COVID-19 diagnosis, providing high-quality databases for various studies. Consequently, Dataset-2 did not The filter used was based on the so-called decision table presented in the work of Ribeiro and Yao [24]. Basically, the filtering clusters the cases into predefined categories, producing the uncertainty. Additionally, some cases were forced to be clumped in the same cluster. In a cyclic process, this strategy gave the Final Dataset prepared for the experiments.present the uncertainty cases (as considered in this paper) necessary to prove the concept proposed in this paper (training an MLP with examples of uncertainty detected by RS).

In order to deal with this drawback, for the sake of proving the concept proposed, it was necessary to filter Dataset-2, generating the uncertainty artificially and creating the so-called Dataset-3 (as shown in Figure 8).

The filter used was based on the so-called decision table presented in the work of Ribeiro and Yao [24]. Basically, the filtering clusters the cases into predefined categories, producing the uncertainty. Additionally, some cases were forced to be clumped in the same cluster. In a cyclic process, this strategy gave the Final Dataset prepared for the experiments.

Decision Table: Ribeiro and Yao [24] used rough sets in their work with LBP descriptors of segmented corn grain images and applied a preprocessing step before building the decision table used in rough set analysis. This step consisted of a mean normalization of each column of the descriptors and data binning. For this, we calculated a δ for each feature of the column as follows:

$$\delta = \frac{x-\mu}{\sigma}100 \tag{4}$$

This step resulted in an intermediate δ-table, with $\delta \in (-100, 100)$, and a function $f(\delta)$ was applied for each value such as:

$$f(\delta) = \begin{cases} 0, & \text{if } \delta < -100 \\ 1, & \text{if } -100 < \delta \leq -80 \\ 2, & \text{if } -80 < \delta \leq -60 \\ \vdots & \\ 16, & \text{if } 60 < \delta \leq 80 \\ 17, & \text{if } 80 < \delta \leq 100 \\ 18, & \text{if } \delta > 100 \end{cases} \tag{5}$$

This resulted in a decision table with values binned in intervals of 20, contained within the $(-100, 100)$ range of the mean normalization. As in Ribeiro and Yao's work [24], different intervals were tested, and an interval of 20 yielded the best results.

5.1.5 Final Dataset – Three Label of Classes

The data preparation aims to get the Final Dataset with cases equally distributed in three classes: COVID-19 cases, Healthy cases, and Uncertainty cases. For this purpose, Dataset-3 was analyzed by the RS method implemented in Python. The libraries selected for this task were TWD [25], which was already used in scientific research [24], and Roughsets-base [26] [27], due to its superior performance and therefore being the main one used.

The results obtained by the *Roughsets-base* library were validated using the *TWD* library. After identifying the boundary elements, the label "uncertainty" was defined for the cases identified by RS as boundary members (as shown in Figure 8).

As seen in Table 3, RS analysis identified the uncertain cases and generated an imbalance between classes (6,800 healthy cases, 5,301 uncertainty, and 8,282 COVID cases). To resolve this situation, the number of uncertainty cases (5,301) was used as the limit, determining a cut in the surplus in the size of the other two groups of cases, providing three balanced groups with 5,301 cases each.

Table 3. Cases Distribution after RS analysis

Classes	Dataset 2
Healthy	6800
Uncertain	5301
COVID-19	8282
Limit for cutting (per classe)	5301

5.2 Multi-Layer Perceptron Training

The scikit-learn [28] Multi-Layer Perceptron (MLP) neural network was used for experiments. It is a feedforward neural network that is widely applied for classification tasks. The Final Dataset was split into a training set and a validation set with a ratio of 80:20, respectively. The training set was used to train the MLP neural network, and the validation set was used to evaluate the performance of the model.

Tuning Hyperparameters: A grid search method was used to tune the hyperparameters of the MLP neural network. This grid search technique tests all possible combinations of hyperparameters and chooses the best ones for the data provided. The hyperparameters that were tuned include the number of hidden layers, the number of neurons in each hidden layer, and the learning rate. The accuracy score was used as the evaluation metric to select the optimal set of hyperparameters. Below is Table 4 with the hyperparameters chosen by grid search.

Table 4. Hyperparameters table

Type	Chosen
learning rate	constant
hidden layer sizes	100
activation	logistic
solver	lbfgs
alpha	0.001
Accuracy before gridsearch:	78.41±0.02%
Accuracy after gridsearch:	88.02±0.02%

5.3 Ethical Considerations

The use of the databases was compliant with the ethical standards and regulations. All databases used anonymized data.

6. Results and Discussion

After this composition, we trained a model with the hyperparameters described in section 5.2 and extracted training and testing metrics from the model. The metrics extracted are described in detail in section 3.1, and the results are in Table 5 below.

Table 5. Evaluation Metrics

Metrics	Micro	Macro
Precision	88.02%	87.95%
Recall	87.61%	87.57%
F1 Score	87.61%	87.51%
Accuracy		87.61%
Cross validation accuracy		88.00±0.02%

As expected, both micro and macro results are very similar since there were no imbalances in the number of images per class, with the micro metrics being slightly more accurate. The precision, recall, and F1 score are also very similar, which means the model can not only produce fewer false positives but also produce fewer false negatives, which is crucial for this application. Below, in Figure 9, is the confusion matrix, and in Figure 10, the ROC curve plot and the AUC for each class.

The model performed very similarly for Normal and COVID-19 cases, represented by the labels ”NORMAL” and ”COVID,”

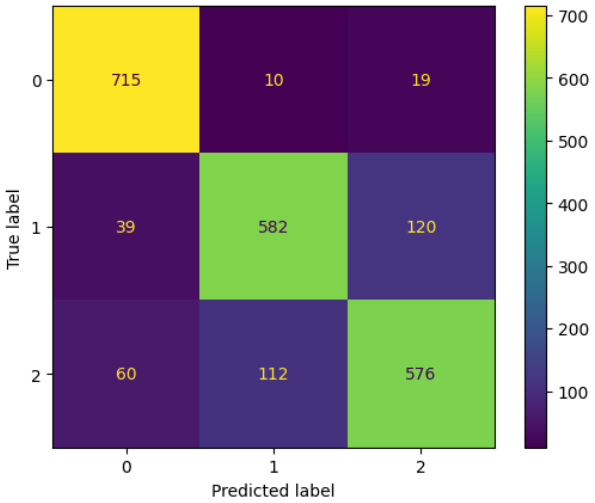


Figure 9. Confusion Matrix

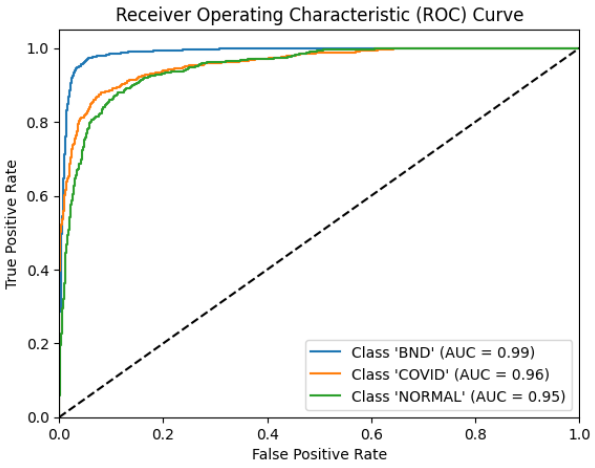


Figure 10. ROC Curve

with an AUC of 0.95 and 0.96, respectively, and had better performance with the uncertain cases represented by the ”BND” label, with an AUC of 0.99, which suggests the model is highly capable of discriminating the uncertain cases from the rest of the cases and accurately classifying them.

These results were similar to the ones obtained by the authors when they tried SVM as the classifier [8]; in this case, the general accuracy was 91.82%. The comparison between SVM and MLP shows that SVM performs better, with an accuracy of 91.82% compared to MLP’s 87.61%. SVM also has higher precision, recall, and F1 scores. The key difference is in their complexity: SVM is not affected by the number of features in the data and can generalize better to new data [29] [30]. On the other hand, MLP’s performance depends on the number of hidden units it uses, which need to be kept small. This makes MLP less effective for high-dimensional data.

Acknowledgements

The authors would like to acknowledge the contributions of the University of the State of Santa Catarina (UDESC) and Research Support Foundation of the State of Santa Catarina (FAPESC) for their financial and structural support in the development of this project.

Author contributions

Kalyl Henings, Vinicius Hansen: Sistematic Review, data preparation and programming, reporting results and documentation; Gilmário Barbosa dos Santos: Academic advising, provisioning resources, supervision, general coordination of the scientific initiation project.

References

[1] WHO Coronavirus (COVID-19) Dashboard. 2023. Disponível em: ⟨https://covid19.who.int⟩.

[2] BARATELLA, E. et al. Severity of lung involvement on chest x-rays in sars-coronavirus-2 infected patients as a possible tool to predict clinical progression: an observational retrospective analysis of the relationship between radiological, clinical, and laboratory data. *J. bras. pneumol.*, v. 46, 2020.

[3] ARSHAD, A. et al. Association of Chest Xray (CXR) findings with outcomes in COVID-19 Patients. *Journal of the Pakistan Medical Association*, v. 72, n. 9, p. 1746–1749, 2022. ISSN 0030-9982. Publisher: Knowledge Bylanes.

[4] BRUNO, M. A. 256 Shades of gray: uncertainty and diagnostic error in radiology. *Diagnosis*, v. 4, n. 3, p. 149–157, set. 2017. ISSN 2194-802X. Publisher: De Gruyter. Disponível em: ⟨https://www.degruyter.com/document/doi/10.1515/dx-2017-0006/html⟩.

[5] KURZ, A. et al. Uncertainty estimation in medical image classification: Systematic review. *JMIR Med Inform.*, v. 10, p. e36427, august 2022.

[6] ABDAR, M. et al. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion*, v. 76, p. 243–297, 2021. ISSN 1566-2535. Disponível em: ⟨https://www.sciencedirect.com/science/article/pii/S1566253521001081⟩.

[7] ZOU, K. et al. A review of uncertainty estimation and its application in medical imaging. *Meta-Radiology*, v. 1, n. 1, p. 100003, 2023. ISSN 2950-1628. Disponível em: ⟨https://www.sciencedirect.com/science/article/pii/S2950162823000036⟩.

[8] HANSEN, V.; HENINGS, K.; SANTOS, G. B. d. Applying FOSS Support Vector Machine and Rough Sets on COVID-19 Cases Triage. In: *Anais do Latin Science 2023*. SBC, 2023. p. 22–27. ISSN: 2160-1348. Disponível em: ⟨https://sol.sbc.org.br/index.php/latinoware/article/view/26077⟩.

[9] HUANG, L. et al. *A review of uncertainty quantification in medical image analysis: probabilistic and non-probabilistic methods*. 2023.

[10] ALIZADEHSANI, R. et al. Handling of uncertainty in medical data using machine learning and probability theory techniques: a review of 30 years (1991-2020). *Ann Oper Res*, v. 21, p. 1–42, 2021.

[11] PAWLAK, Z. Rough set theory and its applications. *Journal of Telecommunications and Information Technology*, nr 3, p. 7–10, 2002. ISSN 1509-4553. Disponível em: ⟨http://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-article-BPS2-0018-0017⟩.

[12] ROUGH Set Theory | An Introduction. 2019. Section: Machine Learning. Disponível em: ⟨https://www.geeksforgeeks.org/rough-set-theory-an-introduction/⟩.

[13] CHEN, Y. et al. Measures of uncertainty for neighborhood rough sets. *Knowledge-Based Systems*, v. 120, p. 226–235, 2017. ISSN 0950-7051. Disponível em: ⟨https://www.sciencedirect.com/science/article/pii/S0950705117300096⟩.

[14] ANACONDA Software Distribution. Anaconda Inc., 2020. Publication Title: Anaconda Documentation Version Number: Vers. 2-2.4.0. Disponível em: ⟨https://docs.anaconda.com/⟩.

[15] RAHMAN, T. et al. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Computers in Biology and Medicine*, v. 132, p. 104319, maio 2021. ISSN 0010-4825. Disponível em: ⟨https://www.sciencedirect.com/science/article/pii/S001048252100113X⟩.

[16] CHOWDHURY, M. E. H. et al. Can AI Help in Screening Viral and COVID-19 Pneumonia? *IEEE Access*, v. 8, p. 132665–132676, 2020. ISSN 2169-3536. Conference Name: IEEE Access.

[17] VAYÁ, M. I. et al. *BIMCV COVID-19+: a large annotated dataset of RX and CT images from COVID-19 patients*. arXiv, 2020. ArXiv:2006.01174 [cs, eess]. Disponível em: ⟨http://arxiv.org/abs/2006.01174⟩.

[18] WANG, Z. et al. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, v. 13, n. 4, p. 600–612, abr. 2004. ISSN 1941-0042. Conference Name: IEEE Transactions on Image Processing.

[19] ALHASHIM, I.; WONKA, P. *High Quality Monocular Depth Estimation via Transfer Learning*. arXiv, 2019. ArXiv:1812.11941 [cs]. Disponível em: ⟨http://arxiv.org/abs/1812.11941⟩.

[20] WALT, S. et al. scikit-image: image processing in Python. *PeerJ*, v. 2, p. e453, jun. 2014. ISSN 2167-8359. Disponível em: ⟨https://doi.org/10.7717/peerj.453⟩.

[21] JT XU S, W. B. K. Efficient data mining for local binary pattern in texture image analysis. *Expert Syst Appl.*, v. 42, n. 9, p. 4529–4539, 2015.

[22] SABRI, N. et al. COVID-19 detection for chest X-ray images using local binary pattern. *International Journal of Emerging Trends in Engineering Research*, p. 78–81, 2020. Disponível em: ⟨https://doi.org/10.30534/ijeter/2020/1281.12020⟩.

[23] THEPADE, S. D.; JADHAV, K. Covid19 Identification from Chest X-Ray Images using Local Binary Patterns with assorted Machine Learning Classifiers. In: *2020 IEEE Bombay Section Signature Conference (IBSSC)*. [S.l.: s.n.], 2020. p. 46–51.

[24] RIBEIRO, S. S.; YAO, J. Toward a Three-Way Image Classification Model: A Case Study on Corn Grain Images. In: *2019 IEEE International Symposium on Multimedia (ISM)*. San Diego, CA, USA: IEEE, 2019. p. 177–1776. ISBN 978-1-72815-606-4. Disponível em: ⟨https://ieeexplore.ieee.org/document/8959001/⟩.

[25] RIBEIRO, S. S. *TWD*. 2023. Original-date: 2021-06-23T17:31:12Z. Disponível em: ⟨https://github.com/profssribeiro/twd⟩.

[26] JANKOWSKI, D. *Roughsets-base: Python library for base Roughsets methods.* Zenodo, 2022. Disponível em: ⟨https://zenodo.org/record/5957474⟩.

[27] JANKOWSKI, D. *darekjk/roughsets-base*. 2023. Original-date: 2022-02-03T08:30:14Z. Disponível em: ⟨https://github.com/darekjk/roughsets-base⟩.

[28] PEDREGOSA, F. et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.

[29] ZANATY, E. A. Support Vector Machines (SVMs) versus Multilayer Perception (MLP) in data classification. *Egyptian Informatics Journal*, v. 13, n. 3, p. 177–183, nov. 2012. ISSN 1110-8665. Disponível em: ⟨https://www.sciencedirect.com/science/article/pii/S1110866512000345⟩.

[30] GHOSH, S.; MONDAL, S.; GHOSH, B. A comparative study of breast cancer detection based on svm and mlp bpn classifier. In: *2014 First International Conference on Automation, Control, Energy and Systems (ACES)*. [S.l.: s.n.], 2014. p. 1–4.