المجلد 16 العدد 4 | مجلة جامعة الانبار للعلوم الاقتصادية والادارية | السنة/2024 م

# Comparing and Evaluating Machine Learning Techniques for Liver-Hydatid-Cyst Classification

**Dr. Saman Hassan Abdulla** saman.abdulla@univsul.edu.iq
**Soran Husen Mohamad** soran.husen1990@gmail.com
Department of Statistics and Information, College of Administration and Economics, University of Sulaimani, Sulaymaniyah, Iraq

# مقارنة وتقييم تقنيات التعلم الآلي لتصنيف الكيس المائي في الكبد

د. سامان حسن عبد الله
جامعة السليمانية، طريق كركوك
السليمانية، العراق

سوران حسين محمد
جامعة السليمانية، طريق كركوك،
السليمانية، العراق

تاريخ استلام البحث 2024/9/22 تاريخ قبول النشر 2024/11/14 تاريخ النشر 2024/ 12/ 30

## Abstract

The liver hydatid cyst is a parasitic infection that is triggered by the presence of E. granulosus, this infection results in the formation of cysts in the liver. It is often preceded by surgery or followed by anti-parasitic treatment. The purpose of the study is to investigate and assess the effectiveness of different machine learning methods in classifying hydatid cases of liver. The dataset used in this investigation was collected from patients at Sulaimani Smart Hospital during a period in 2024. It comprises records from patients in the 3rd party. The algorithms employed for contrast of Bagging (Bg),JRip (JR), OneR (OR) and partial decision tree (PART) each have their own advantages and processes associated with data classification. The performance of these algorithms is judged using multiple criteria that primarily focus on the confusion matrix, including Accuracy (the number of correct classifications and the percentage of correct classifications), and Sensitivity. The analysis of the data indicates that Bagging (Bg) has the greatest capacity for classification, with a maximum accuracy of 74.8 percent, and has been successfully applied to ninety two cases, which are effectively classified, and 31 instances, which are incorrectly classified. Other classifiers have a lower capacity for classification. Bagging is additionally notable for its high sensitivity of 96.74% among all classifiers.

***Keywords:*** *Locally Weighted Learning, Partial Decision Tree, Receiver Operating Characteristic*

**المستخلص**

الكيس العداري في الكبد هو عدوى طفيلية تنجم عن وجود بكتيريا الإشريكية الحبيبية، وتؤدي هذه العدوى إلى تكوين كيسات في الكبد. وغالبًا ما يسبقه إجراء عملية جراحية أو يتبعه علاج مضاد للطفيليات. الغرض من هذه الدراسة هو دراسة وتقييم فعالية طرق التعلم الآلي المختلفة في تصنيف حالات الكبد العدارية." تم جمع مجموعة البيانات المستخدمة في هذا البحث من مرضى في مستشفى السليمانية الذكي خلال فترة في عام 2024. وهي تشتمل على سجلات من المرضى في الطرف الثالث. تتمتع كل من الخوارزميات المستخدمة لتباين التعبئة (Bg) و JRip (JR) وOneR (OR) وشجرة القرار الجزئي (PART) بمزايا وعمليات خاصة مرتبطة بتصنيف البيانات. يتم الحكم على أداء هذه الخوارزميات باستخدام معايير متعددة تركز بشكل أساسي على مصفوفة الارتباك، بما في ذلك الدقة (عدد التصنيفات الصحيحة والنسبة المئوية للتصنيفات الصحيحة)، والحساسية. ويشير تحليل البيانات إلى أن التعبئة (Bg) تتمتع بأكبر قدرة على التصنيف، مع دقة قصوى تبلغ 74.8 بالمئة، وقد تم تطبيقها بنجاح على اثنتين وتسعين حالة تم تصنيفها بشكل فعال، و31 حالة تم تصنيفها بشكل غير صحيح. تتمتع المصنفات الأخرى بقدرة أقل على التصنيف. تتميز عملية التعبئة أيضًا بحساسيتها العالية التي تصل إلى 96.74% بين جميع المصنفات.


***الكلمات المفتاحية:*** *التعلم المحلي الموزون، شجرة القرار الجزئية، منحنى الخصائص التشغيلية المستقبلية*

## 1. Introduction

The categorization of medical files has a significant role in the diagnosis and treatment of diseases, especially in regards to liver disease. Among the conditions, liver hydatid cyst (LHC) poses a significant challenge due to its complexity and multiple symptoms. Precise classification and analysis are crucial to the administration of this condition with a remarkable solution (Abdulameer & Abdullah, 2018: 291). Recent advances in machine learning (ML) have augmented a variety of methods that can enhance the accuracy and effectiveness of scientific information classification (Ağbulut, Gürel, & Biçen, 2021: 4). This investigation focuses on evaluating four effective methods of machine learning: Bagging (Bg),JRip (JR), OneR (OR), and partial decision tree (PART) for the classification of data regarding liver-hydatid-cyst. Machine learning is a subset of synthetic intelligence that involves computers that analyze patterns from data and make predictions or judgments based on the information (Ahmed, Mohamad, & Karim, 2023: 267), (Ahmed, Hamdin, & Mohamad, 2023: 296), Alam, Mehmood, & Katib, 2020: 145). Demonstrated that ensemble machine learning methods, particularly Bagging and Boosting, significantly enhance the accuracy of face recognition systems, achieving correct classification rates of up to 99%, especially when utilizing Random Forest decision trees. Similarly, highlighted the critical need for comparing various filtering techniques within recommendation systems. Their research explored algorithms such as BayesNet, Decision Table, Logistic Regression, k-NN, JRip, LibSVM, and Random Forest, employing evaluation metrics such as Kappa Statistic and Accuracy (Doulah, 2019: 4). Notably, Random Forest achieved an impressive accuracy of 99.39% on the MovieLens dataset using the WEKA tool. Furthermore, (Kumar, Viinikainen, & Hamalainen, 2017: 266) studied the integration of outputs from five supervised machine learning algorithms, including Random Forest (RF), PART, JRIP, J.48, and Ridor. The evaluation results indicated that the proposed model effectively detected both known and unknown threats, achieving an accuracy rate of 98.2%. This study of objectives that will reveal the effectiveness of these methods in differentiating between one-of-a-kind training of hydatid-cyst data that is valuable, pertinent to scientific professionals and researchers. Algorithm-based machine learning is essential to a variety of fields: in healthcare for the prognosis of illnesses and the prediction of results, in finance for the detection of fraud and the assessment of hazards, in marketing and advertising for the segmentation of clients and the personalization of recommendations, and in technological expertise for the recognition of speech, the classification of photographs, and the sustaining of vehicles (Omer, Faraj, & Mohamad, 2023: 26).

## 2. The target of this study

The objective of this analyze about is to compare and evaluate a variety machine learning to determine their effectiveness in precisely classifying liver hydatid cyst cases. This consists of examining the universal overall performance of extraordinary algorithms to pick out the most reliable and environment friendly method for diagnosing the condition. The dataset for this study was gathered from patients at Sulaimani Smart Hospital in 2024, comprising archives from 123 patients. This research about used four one of a kind pc getting to be aware of methods such as Bagging (Bg), JRip (JR), OneR (OR) and Partial Decision Tree (PART).

## 3. Materials and Methods

There are a variety of classification algorithms that has Bagging (Bg), JRip (JR), OneR (OR) and Partial Decision Tree (PART) classification applications and investigated in Liver-Hydatid-Cyst data. The working process is proven in Figure 1. We selected classification algorithm to Rephrase the most suitable one for predicting Liver-Hydatid-Cyst data.

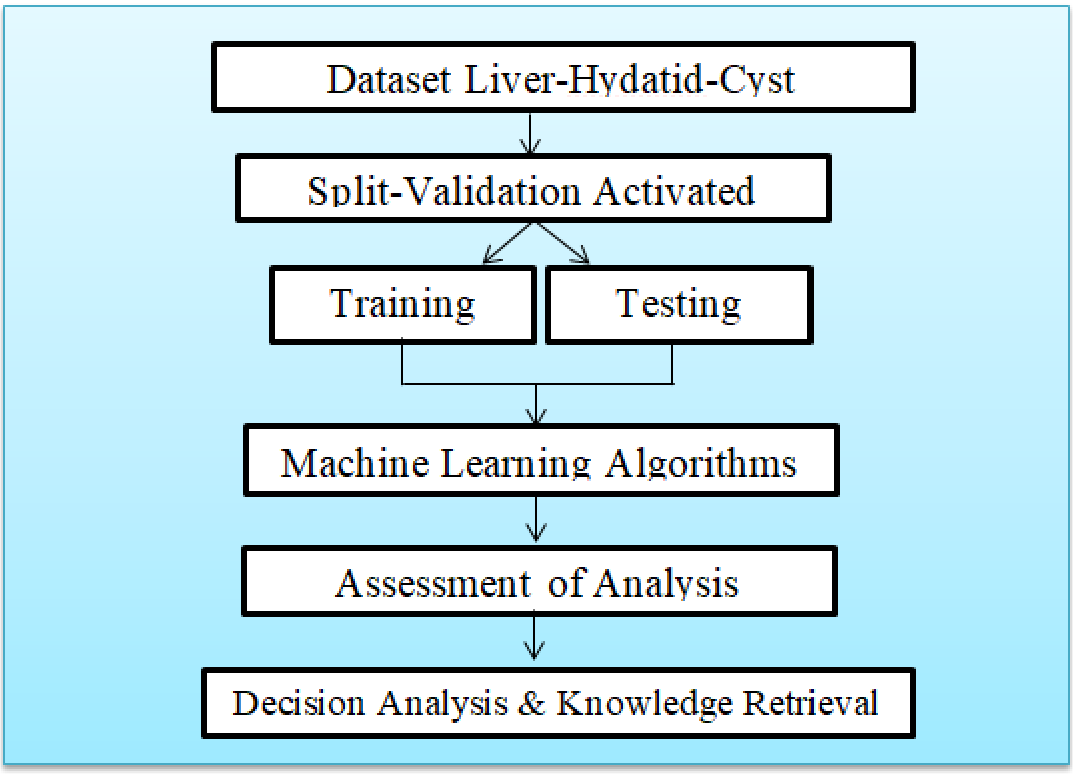


Figure (1): working process

In this study, the classification of Liver-Hydatid-Cyst data will be conducted using four widely utilized machine learning algorithms, including:

## 3.1 Bagging (Bg)

Bagging, additionally recognized as Bootstrap Aggregating, is a robust ensemble mastering technique that improves mannequin steadiness and classification accuracy. The technique makes use of choice sampling to resample the schooling facts and create a couple of subsets. After that, a one of a sort classifier is trained on each subset, and to achieve a last decision, the predictions of these classifiers are combined, generally via voting (Bauer & Kohavi, 1999: 110). By the use of this method, the effects of noisy data or large variance in man or girl classifiers are lessened and variance is decreased. Bagging is a really useful tactic for acquiring increased dependable and consistent classification results in machine learning getting to know on account that it enhances standard overall performance and generalisation via ability of combining the outputs of quite a few models (Yaman, Rattay, & Subasi, 2021: 205).

## 3.2 JRip (JR)

The JRip classifier is a machine learning algorithm used for rule-based classification (Alam, Ubaid, Sohail, Nadeem, Hussain, & Siddiqui, 2021: 214). It is an implementation of the RIPPER (Repeated Incremental Pruning to Produce Error Reduction) algorithm. JRip constructs a set of regulations for classifying cases in a dataset via capability of iteratively refining and pruning them to improve accuracy (Chauhan, Kumar, Pundir, & Pilli, 2013: 42). The ensuing regulations are both interpretable and effective, making JRip a treasured tool for responsibilities the place grasp the decision-making machine is crucial. It balances rule complexity and accuracy, supplying a sturdy technique for dealing with more than a few archives kinds and classifications, and is in particular useful for generating fashions that are trouble-free to interpret (Velmurugan & Anuradha, 2016: 59).

## 3.3 OneR (OR)

A machine learning algorithm used for rule-based classification is called the OneR classifier. It streamlines the classification process by providing the advantageous resource of creating a single rule that, when applied alone, offers the best known daily performance based only on the attribute (Ali, Abdullah, & Mohamad, 2023: 144). OneR assesses each characteristic separately to determine which one best predicts the target kind with the fewest errors. After that, it creates a rule that is mainly based on this attribute. Even though OneR is straightforward, it can be a great tool for providing a baseline mannequin that is accessible and for providing insights into which elements are most useful for classification tasks (Singh, 2009: 482).

## 3.4 Partial Decision Tree (PART)

he PART classifier is a machine learning algorithm designed for rule-based classification. It combines elements of decision tree learning and rule-based approaches to create an effective and interpretable model. PART constructs a decision tree to identify potential decision paths, then translates these paths into a set of rules (Berger, Merkl, & Dittenbach, 2006: 1107). Each rule represents a specific condition or combination of conditions derived from the data, which can be used to classify new instances. This method allows PART to balance the complexity of decision trees with the clarity of rule-based systems, making it a valuable tool for tasks that require both accuracy and interpretability (Ozturk Kiyak, Tuysuzoglu, & Birant, 2023: 8).

## 4. Performance Evaluation

1. A confusion matrix is a valuable tool for assessing the performance of a classifier, as demonstrated in Table 1. It outlines the count of correctly identified negative instances, termed true negatives (TN), alongside the accurately identified positive cases, referred to as true positives (TP). False positives (FP) indicate negative instances that have been erroneously classified as positive, while false negatives (FN) signify positive cases that have been incorrectly categorized as negative (Jain, 2023: 1190). This matrix provides a comprehensive analysis of the classifier's accuracy and error distribution, highlighting both its advantages and limitations.

Table (1) Confusion Matrix

| **Actual** | **Predicted** | |
|---|---|---|
| | Positive | Negative |
| Positive | TP | FN |
| Negative | FP | TN |

A thorough precis of the classifier contrast metrics is provided in Table 2, which also includes the following: F-Measure, Classification Accuracy, Sensitivity (True Positive Rate), Specificity, False Positive Rate, Precision, Recall, and Area Under the Receiver Operating Characteristic Curve (ROC Area). Classification Accuracy counts the proportion of cases that are efficiently classified; Sensitivity analyses how properly the mannequin detects high quality instances; and Specificity analyses how properly it detects poor instances. Recall gauges how nicely the model captured every actual positive, whilst precision indicates the proportion of true positives amongst positive predictions. Precision and Recall are blended in the F-Measure to grant a balanced performance metric. The ROC Area shows how nicely the mannequin can distinguish between classes, and the MCC gives an average overall performance measure by using accounting for all of the confusion matrix's elements. When these metrics are combined, a thorough evaluation of the classifier's effectiveness is produced, demonstrating each its prediction accuracy and efficiency (Yousefi & Poornajaf, 2023: 2).

Table (2) Detailed Accuracy By Classes

| **Tools** | **Statistic** |
|---|---|
| Accuracy | $\frac{TP+TN}{TP+TN+FP+FN} \times 100$ |
| Sensitivity (Recall or True Positive Rate) | $\frac{TP}{TP+FN} \times 100$ |
| Specificity (True Negative Rate) | $\frac{TN}{TN+FP} \times 100$ |
| TP Rate (True Positive Rate) | $\frac{TP}{TP+FN} \times 100$ |
| FP Rate (False Positive Rate) | $\frac{FP}{FP+TN} \times 100$ |
| Precision (Positive Predictive Value) | $\frac{TP}{TP+FP} \times 100$ |

| | |
|---|---|
| Recall (Sensitivity or True Positive Rate) | $\frac{TP}{TP+FN} \times 100$ |
| F-Measure (F1 Score) | $\frac{2.(Precision . Recall)}{Precision+ Recall} \times 100$ |
| Matthews Correlation Coefficient (MCC) | $\frac{TP.TN-FP.FN}{\sqrt{(TP+FP)(TP+FN)(TN+Fp)+(TN+FN)}} \times 100$ |
| Receiver Operating Characteristic (ROC) | $\frac{TP+TN}{TP+FN+TN+FP} \times 100$ |

## 5. Data Analysis

This paper focuses on comparing and evaluating several machine learning techniques tailored for the classification of Liver-Hydatid-Cyst data. Specifically, it examines Bagging (Bg), JRip (JR), OneR (OR), and Partial Decision Tree (PART). The 123 patient files that made up the dataset for this study was gathered in 2024 from patients at Sulaimani Smart Hospital. The following variables had been existing in the facts set: The explanatory variables are represented with the aid of the first group: Age, gender, region of residence, BMI, HT, DM, hypothyroidism, and preoperative cyst variety. The response variable is indicated Size of Cyst and has the following code: Based on size, the cysts are separated into two most important groups. Less than 10 cm in diameter cysts are covered in the first group, "Small to Medium." These cysts may no longer purpose any signs at all, or they might also cause moderate to moderate symptoms like strain or discomfort. The second group, "Large," covers cysts that are 10 cm or larger. Larger cysts are extra in all likelihood to produce serious symptoms like pain, jaundice, or issues like rupture or secondary infections. Machine learning algorithms were used in Weka (Hall, Frank, Holmes, Pfahringer, Reutemann, & Witten, 2009: 115).

### 5.1 Confusion Matrix for Machine learning algorithms

The performance of four machine learning algorithms Bagging (Bg), JRip (JR), OneR (OR), and Partial Decision Tree (PART) in classifying data into two classes, a and b, is compared in the confusion matrix that is provided. Out of 123 instances, the model accurately predicted class b for Bagging (Bg) 89 times and class a 3 times. On the other hand, it incorrectly identified class b as class a three times and class an as class b 28 times. As a result, there are 117 accurate predictions and 6 incorrect ones overall.

JRip (JR) displayed somewhat different results: 4 class B predictions were correct, and 88 class a predictions were correct. 26 false positives, 29 false positives, and 117 accurate predictions were produced overall 29 times class a was classified as class b and twice class b as class a. With class a being correctly predicted 85 times and class b only seven times, OneR (OR) suggested a lower accuracy for class a. Four times, class b was mistakenly assigned to class a, and twenty-seven times, class was mistakenly assigned to class b. Of the predictions made, 112 were accurate and 11 were incorrect.

The partial decision tree (PART), which efficaciously expected classification a seventy seven instances and classification b 15 times, confirmed the lowest accuracy for that class. There had been 18 misclassifications out of 105 correct predictions, with 28 cases of category a being expected as class b and three cases of class b being estimated as type a. When it came to normal performance and misclassification rates, Bagging and JRip performed the best, whilst OneR and PART showed greater prices of misclassification and the lowest usual accuracy.

Table (3) Confusion matrix for machine learning algorithms

| Bagging (Bg) | | | | JRip (JR) | | | |
|---|---|---|---|---|---|---|---|
| Predicted Class | | | | Predicted Class | | | |
| Actual Class | a | b | Total | Actual Class | a | b | Total |
| A | 89 | 3 | 92 | a | 88 | 4 | 92 |
| B | 28 | 3 | 31 | b | 29 | 2 | 31 |
| Total | 117 | 6 | 123 | Total | 117 | 6 | 123 |
| OneR (OR) | | | | Partial Decision Tree (PART) | | | |
| Predicted Class | | | | Predicted Class | | | |
| Actual Class | a | b | Total | Actual Class | a | b | Total |
| A | 85 | 7 | 92 | a | 77 | 15 | 92 |
| B | 27 | 4 | 31 | b | 28 | 3 | 31 |
| Total | 112 | 11 | 123 | Total | 105 | 18 | 123 |

.

## 5.2 Classification Accuracy, Sensitivity and Specificity of Proposed

The performance of the four classifiers varies with recognize to sensitivity, specificity, accuracy, and correct/incorrect classifications. They are Partial Decision Tree (PART), OneR (OR), JRip (JR), and Bagging (Bg). Bagging (Bg), which attains the highest sensitivity of 96.74%, demonstrates the potential to discover wonderful cases with excessive accuracy. Conversely, its 9.68% specificity is remarkably low and suggests challenges in precisely classifying terrible cases. A 25.20% improper classification rate used to be obtained from the ninety two cases that have been correctly labeled and the 31 instances that were incorrectly classified. This suggests that bagging accuracy is 74.80% overall.

Similar to bagging, JRip (JR) has a high sensitivity of 95.65% but an even lower specificity of 6.45%. With 90 cases correctly classified and 33 incorrectly classified, or a 26.83% incorrect classification rate, its overall accuracy is slightly lower at 73.17%. OneR (OR) outperforms the other classifiers with a higher specificity of 12.90% but a lower sensitivity of 92.39% when compared to Bagging and JRip. With 89 cases correctly classified and 34 incorrectly classified, it has an overall accuracy of 72.36% and an incorrect classification rate of 27.64%.

Partial decision timber (PART) have the perfect specificity (9.68%) and the lowest sensitivity (83.70%), a whole lot like bagging. It has the lowest universal accuracy (65.04%), with 34.96% of cases incorrectly classified and 80 cases effectively categorized out of forty three errors. Based on the share of correctly classified instances, Bagging (Bg) surpasses all other classifiers with an accuracy of 74.80%. Bagging is the satisfactory choice due to the fact of its excessive sensitivity and typical accuracy, even although its specificity is low, specifically if the principal intention is to correctly classify most instances.

Table (4) : The Classification Accuracy, Sensitivity and Specificity of Classifier

| Classifier | Sensitivity % | Specificity % | Accuracy % | N. Correctly | Correctly Classified % | N. Incorrectly | Incorrectly Classified % |
|---|---|---|---|---|---|---|---|
| Bagging (Bg) | 96.74 | 9.68 | 74.80 | 92 | 74.80 | 31 | 25.20 |
| JRip (JR) | 95.65 | 6.45 | 73.17 | 90 | 73.17 | 33 | 26.83 |
| OneR (OR) | 92.39 | 12.90 | 72.36 | 89 | 72.36 | 34 | 27.64 |
| Partial Decision Tree (PART) | 83.70 | 9.68 | 65.04 | 80 | 65.04 | 43 | 34.96 |

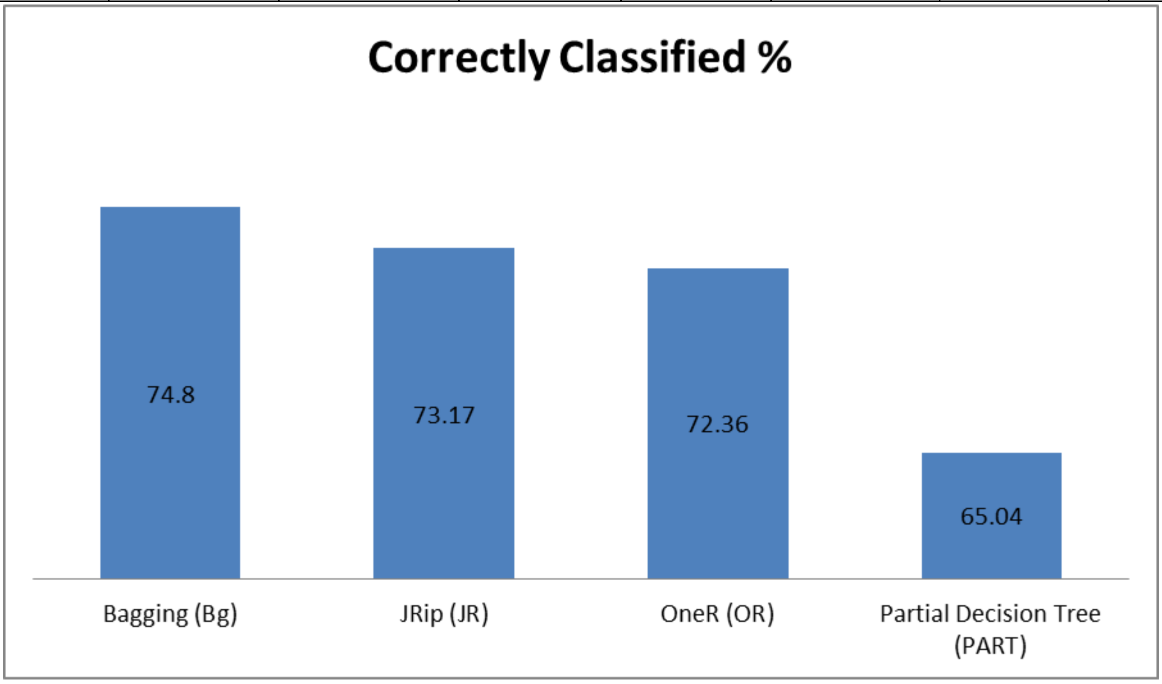


Figure (2) Illustrate the difference in correctly classified % of the proposed models.

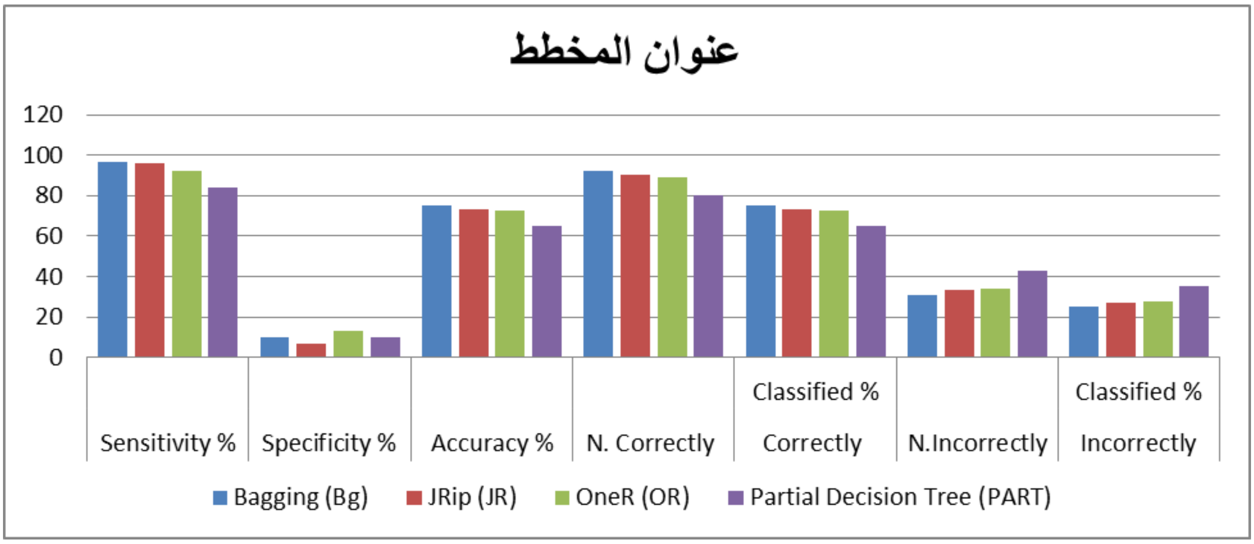


Figure (3) Illustrate the detailed accuracy of the proposed models

## 5.3 Calculation Detailed Performance Metrics

Each classifier's complete set of performance metrics Their classification performance for classes 'a' and 'b' is demonstrated by JRip (JR), OneR (OR), Bagging (Bg), and Partial Decision Tree (PART). True Positive (TP) Rate for class 'a' is 0.967, indicating that bagging performs well in identifying positive cases. Both its high False Positive (FP) Rate of 0.903 and low TP Rate of 0.097 for class "b" suggest that it struggles with classifying negatives. By looking at its overall weighted metrics, its TP Rate is 0.748 and its ROC Area is 0.437. JRip also does a great job of identifying positives; for class "a," its TP Rate is 0.957. However, it also has a high FP Rate of 0.935 and only moderately. OneR gives better specificity for classification 'b' (TP Rate: 0.129) and a greater balanced performance for category 'a' (TP Rate: 0.924). With the highest ROC Area of 0.526 and a TP Rate of 0.724, its normal metrics display regular overall performance in each classes. Due to its low TP Rate of 0.097 for category 'b', Partial Decision Tree (PART) has bother with negatives. The category 'a' has the lowest TP Rate, which is 0.837. A 0.65 TP Rate and a 0.511 ROC Area are proven by way of the standard weighted metrics of PART. Based on its most appropriate aggregate of precision, recall, and ROC Area, OneR (OR) is the most reliable classifier for balanced performance..

Table (5) Detailed Performance Metrics by Class for Classifiers

| Detailed Accuracy By Classes | | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area |
|---|---|---|---|---|---|---|---|---|
| Bagging (Bg) | a | 0.967 | 0.903 | 0.761 | 0.967 | 0.852 | 0.129 | 0.437 |
| | b | 0.097 | 0.033 | 0.5 | 0.097 | 0.162 | 0.129 | 0.437 |
| | Weighted Avg. | 0.748 | 0.684 | 0.695 | 0.748 | 0.678 | 0.129 | 0.437 |
| JRip (JR) | a | 0.957 | 0.935 | 0.752 | 0.957 | 0.842 | 0.042 | 0.505 |
| | b | 0.065 | 0.043 | 0.333 | 0.065 | 0.108 | 0.042 | 0.505 |
| | Weighted Avg. | 0.732 | 0.711 | 0.647 | 0.732 | 0.657 | 0.042 | 0.505 |
| OneR (OR) | a | 0.924 | 0.871 | 0.759 | 0.924 | 0.833 | 0.081 | 0.526 |
| | b | 0.129 | 0.076 | 0.364 | 0.129 | 0.19 | 0.081 | 0.526 |
| | Weighted Avg. | 0.724 | 0.671 | 0.659 | 0.724 | 0.671 | 0.081 | 0.526 |
| Partial Decision Tree (PART) | a | 0.837 | 0.903 | 0.733 | 0.837 | 0.782 | -0.081 | 0.511 |
| | b | 0.097 | 0.163 | 0.167 | 0.097 | 0.122 | -0.081 | 0.511 |
| | Weighted Avg. | 0.65 | 0.717 | 0.591 | 0.65 | 0.616 | -0.081 | 0.511 |

## 6. Discussion

Significant variations in the overall performance of extraordinary computer mastering algorithms for Liver-Hydatid-Cyst information classification are published by using analysis. After evaluation, Bagging (Bg) is determined to be the most environment friendly algorithm with an remarkable accuracy of 74.80%. Its overall performance metrics, which show that it effectively labeled ninety two out of 123 cases, display its sturdy classification capability and high accuracy. With a sensitivity of 96.74%, bagging additionally had the best possible performance, demonstrating its potent capacity to accurately become aware of positive cases. But at 9.68%, its specificity is notably low, indicating that though Bagging is appropriate at spotting nice cases, it has trouble accurately figuring out terrible ones. With a sensitivity of 95.65% and an accuracy of 73.17%, JRip (JR) is closely behind. It is even less specific at 6.45% than Bagging, despite having comparable strengths in positive identification. Compared to Bagging and JRip, OneR (OR) provides a more balanced performance with a sensitivity of 92.39% and a specificity of 12.90%, making it more trustworthy for classifying both positive and negative cases. Its overall accuracy of 72.36% is marginally less, though. With the lowest accuracy of 65.04%, Partial Decision Tree (PART) performs the worst across all metrics, demonstrating its relative inefficiency in identifying both positives and negatives. These conclusions are further supported with the aid of the complete overall performance metrics, which show that OneR gives a greater balanced method whilst Bagging and JRip excel in sensitivity but have boundaries in specificity. then, the Bagging (Bg) is the best-performing classifier for the dataset in this study due to the fact of its extremely good sensitivity and accuracy.Overall, these results show that though Bagging (Bg) is the most dependable and accurate classifier, other algorithms also work well, each with unique advantages in more than a few areas of classification performance.

## 7. Conclusions and recommendations

### 7.1 Conclusion

With the fantastic accuracy (74.80%) and sensitivity (96.74%), the bagging (Bg) algorithm is the most profitable in classifying liver-hydrostatin-cyst data. Its accuracy in identifying advantageous cases makes it the preferred choice even although its specificity is low. When in contrast to Bagging, JRip (JR) well-knownshows a slightly lower accuracy rate (73.17%) however nonetheless shows excessive sensitivity (95.65%). Given its extraordinarily low specificity, correctly classifying negative instances seems to be challenging. In scenarios where a balanced approach to advantageous and terrible case classification is required, OneR (OR) gives a dependable desire with a sensitivity of 92.39% and specificity of 12.90%, providing a more balanced classification performance. The partial choice tree (PART) performs the worst overall, outperforming other strategies in terms of sensitivity (83.70%) and accuracy (65.04%).

These data are analysed, and Bagging (Bg) shows better classification performance with 74.80% accuracy, correctly classifying 92 cases and incorrectly classifying 31 cases. Given that bagging's sensitivity is 96.74%, this suggests that it is especially good at correctly identifying positive cases. The algorithm's overall accuracy and sensitivity indicate that it is the best choice for applications where the accurate detection of positive instances is the most important factor, despite its relatively low specificity of 9.68%, which results in a higher rate of misclassified negative cases. Because missing a positive case could have serious consequences in medical diagnostics, Bagging's high sensitivity suggests that it is adept at detecting the presence of liver hydatid cysts.

## 7.2 Recommendations

1. Give Bagging (Bg) a higher priority in functions the place precise high quality case identification and high sensitivity are essential, even although its specificity is lower. When minimising false negatives is the aim, it works best.
2. For conditions requiring a more balanced performance between sensitivity and specificity, OneR (OR) is an achievable option. It gives a reliable method when managing both false positives and false negatives.
3. Because JRip (JR) has a very low specificity, it may not be terrific for applications where precise poor case classification is required. Therefore, reevaluate its use in conditions where higher specificity is required.
4. Partial Decision Tree (PART) currently well-known shows the least effectiveness among the evaluated algorithms; therefore, if its decrease accuracy and sensitivity are deemed insufficient, look into other methods or improvements.

## References


1. Abdulameer, M. M. H., & Abdullah, M. Z. (2018). Toward Generating a new Video Education Lectures Dataset and Performance Comparison with Various Machine Learning Algorithms 22(12):279-298.
2. Ağbulut, Ü., Gürel, A. E., & Biçen, Y. (2021). Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. Renewable and Sustainable Energy Reviews, 135, 110114.
3. Ahmed, D. H., Mohamad, S. H., & Karim, R. H. R. (2023). Using Single Exponential Smoothing Model and Grey Model to Forecast Corn Production in Iraq during the period (2022-2030). University of Kirkuk Journal For Administrative and Economic Science, 13(3).
4. Ahmed, N. M., Hamdin, A. O., & Mohamad, S. H. (2023). Copula Probability Model for Determining Dependency and Generations. University of Kirkuk Journal for Administrative and Economic Science, 13(2).
5. Alam, F., Mehmood, R., & Katib, I. (2020). Comparison of decision trees and deep learning for object classification in autonomous driving. Smart Infrastructure and Applications: Foundations for Smarter Cities and Societies, 135-158.
6. Alam, M. T., Ubaid, S., Sohail, S. S., Nadeem, M., Hussain, S., & Siddiqui, J. (2021). Comparative analysis of machine learning based filtering techniques using MovieLens dataset. Procedia Computer Science, 194, 210-217.
7. Ali, F. K. H., Abdullah, S. A., & Mohamad, S. H. (2023). Relationship Between Socio-Demographic Characteristics and Food Labeling Consumption in Sulaimani City by Using Chi-Square Test. Al-Qadisiyah Journal For Agriculture Sciences, 13(1).
8. Bauer, E., & Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. Machine learning, 36, 105-139.
9. Berger, H., Merkl, D., & Dittenbach, M. (2006, April). Exploiting partial decision trees for feature subset selection in e-mail categorization. In Proceedings of the 2006 ACM symposium on Applied computing (pp. 1105-1109).
10. Chauhan, H., Kumar, V., Pundir, S., & Pilli, E. S. (2013, August). A comparative study of classification techniques for intrusion detection. In 2013 International Symposium on Computational and Business Intelligence (pp. 40-43).
11. Doulah, M. (2019). Application of machine learning algorithms in bioinformatics. J Proteom Bioinform, 3(1), 1-11.
12. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 11(1), 10-18.
13. Jain, R. (2023). Comparative Study of Machine Learning Algorithms for Document Classification. International Journal of Computer Sciences and Engineering IJCSE, 7(6):1189-1191.
14. Kumar, S., Viinikainen, A., & Hamalainen, T. (2017, December). Evaluation of ensemble machine learning methods in mobile threat detection. In 2017 12th

International Conference for Internet Technology and Secured Transactions (ICITST) (pp. 261-268).
15. Omer, A., Faraj, S. M., & Mohamad, S. H. (2023). An application of two classification methods: hierarchical clustering and factor analysis to the plays PUBG. Iraqi Journal of Statistical Sciences, 20(1), 25-42.
16. Ozturk Kiyak, E., Tuysuzoglu, G., & Birant, D. (2023). Partial Decision Tree Forest: A Machine Learning Model for the Geosciences. Minerals, 13(6).
17. Singh, P. (2009). Comparing the effectiveness of machine learning algorithms for defect prediction. International Journal of Information Technology and Knowledge Management, 2(2), 481-483.
18. Velmurugan, T., & Anuradha, C. (2016). Performance evaluation of feature selection algorithms in educational data mining. Performance Evaluation, 5(02).
19. Yaman, M. A., Rattay, F., & Subasi, A. (2021). Comparison of bagging and boosting ensemble machine learning methods for face recognition. Procedia Computer Science, 194, 202-209.
20. Yousefi, S., & Poornajaf, M. (2023). Comparison of result of machine learning algorithms in predicting heart disease. Frontiers in Health Informatics, 12.