









ORIGINAL

Comparative Analysis of Advanced Data Mining Methods for Enhancing Medical Diagnosis and Prognosis

Análisis Comparativo de Métodos Avanzados de Minería de Datos para Mejorar el Diagnóstico y Pronóstico Médico

Mohammad Al-Batah¹  , Mowafaq Salem Alzboon¹  , Muhyeeddin Alqaraleh²  , Fawaz Ahmad Alzaghoul³ 

¹Jadara University, Faculty of Information Technology. Irbid, Jordan.

²Zarqa University, Faculty of Information Technology. Zarqa, Jordan.

³Applied Science Private University, Department of Computer Science, Amman Jordan.

Cite as: Al-Batah M, Salem Alzboon M, Alqaraleh M, Ahmad Alzaghoul F. Comparative Analysis of Advanced Data Mining Methods for Enhancing Medical Diagnosis and Prognosis. Data and Metadata. 2024; 3:.465. <https://doi.org/10.56294/dm2024.465>


Submitted: 22-02-2024

Revised: 12-06-2024

Accepted: 02-11-2024

Published: 03-11-2024

Editor: Adrián Alejandro Vitón-Castillo 

Corresponding author: Mohammad Al-Batah 

ABSTRACT

Accurate and early diagnosis, coupled with precise prognosis, is critical for improving patient outcomes in various medical conditions. This paper focuses on leveraging advanced data mining techniques to address two key medical challenges: diagnosis and prognosis. Diagnosis involves differentiating between benign and malignant conditions, while prognosis aims to predict the likelihood of recurrence after treatment. Despite significant advances in medical imaging and clinical data collection, achieving high accuracy in both diagnosis and prognosis remains a challenge. This study provides a comprehensive review of state-of-the-art machine learning and data mining techniques used for medical diagnosis and prognosis, including Neural Networks, K-Nearest Neighbors (KNN), Naïve Bayes, Logistic Regression, Decision Trees, and Support Vector Machines (SVM). These methods are evaluated on their ability to process large, complex datasets and produce actionable insights for medical practitioners. We conducted a thorough comparative analysis based on key performance metrics such as accuracy, Area Under the Curve (AUC), precision, recall, and specificity. Our findings reveal that Neural Networks consistently outperform other techniques in terms of diagnostic accuracy and predictive capacity, demonstrating their robustness in handling high-dimensional and nonlinear medical data. This research underscores the potential of advanced machine learning algorithms in revolutionizing early diagnosis and effective prognosis, thus facilitating more personalized treatment plans and improved healthcare outcomes.

Keywords: Data Mining; Machine Learning; Medical Diagnosis; Prognosis Prediction; Neural Networks.

RESUMEN

Un diagnóstico preciso y temprano, junto con un pronóstico exacto, es fundamental para mejorar los resultados de los pacientes en diversas condiciones médicas. Este artículo se centra en aprovechar técnicas avanzadas de minería de datos para abordar dos desafíos médicos clave: el diagnóstico y el pronóstico. El diagnóstico implica diferenciar entre condiciones benignas y malignas, mientras que el pronóstico busca predecir la probabilidad de recurrencia después del tratamiento. A pesar de los avances significativos en la obtención de imágenes médicas y la recopilación de datos clínicos, lograr una alta precisión en el diagnóstico y el pronóstico sigue siendo un desafío. Este estudio ofrece una revisión exhaustiva de las técnicas de aprendizaje automático y minería de datos más avanzadas utilizadas para el diagnóstico y pronóstico médico, incluyendo Redes Neuronales, K-Nearest Neighbors (KNN), Naïve Bayes, Regresión Logística, Árboles

de Decisión y Máquinas de Vectores de Soporte (SVM). Estos métodos son evaluados por su capacidad para procesar grandes y complejos conjuntos de datos y generar información útil para los profesionales médicos. Realizamos un análisis comparativo detallado basado en métricas clave de rendimiento, como la precisión, el Área Bajo la Curva (AUC), la precisión, el recall y la especificidad. Nuestros hallazgos revelan que las Redes Neuronales superan consistentemente a otras técnicas en términos de precisión diagnóstica y capacidad predictiva, demostrando su solidez en el manejo de datos médicos de alta dimensión y no lineales. Esta investigación destaca el potencial de los algoritmos avanzados de aprendizaje automático para revolucionar el diagnóstico temprano y el pronóstico efectivo, facilitando así planes de tratamiento más personalizados y mejores resultados en la atención médica.

Palabras clave: Minería de Datos; Aprendizaje Automático; Diagnóstico Médico; Predicción de Pronóstico; Redes Neuronales.

INTRODUCTION

Breast cancer is the most commonly diagnosed cancer and the driving cause of cancer passing in women around the globe. Within the US alone, there will be almost 268 600 recently analyzed female breast cancer cases and 41 760 passings for 2019. Early discovery is pivotal; it can diminish the breast cancer mortality rate. To diagnose breast cancer at its different stages; different imaging methods may well be used such as mammography, attractive reverberation imaging (MRI), positron outflow tomography (PET), Computed tomography (CT), and single-photon emanation computed tomography (SPECT).⁽¹⁾

Breast cancer is a tumor, either benign or malignant, that develops when cells in the breast divide and multiply uncontrollably.⁽²⁾ Due to a few risk factors that raise a woman's likelihood of acquiring breast cancer, scientists have attempted to determine the precise cause of the disease. Among the variables being taken into account for breast cancer are age, genetic risk, and family history.⁽³⁾

The goal of the classification algorithms is to construct a model from a set of training data whose target class labels are known and then this model is used to classify unseen instances. The classification of Breast Cancer data can be useful to predict the outcome of some diseases or discover the genetic behavior of tumors.⁽⁴⁾

There are two categories of breast cancer treatments: systemic and local. While chemotherapy and hormone therapy are examples of systematic therapies, surgery, and radiation therapy are instances of localized treatments.⁽⁵⁾ Both therapies are used in combination, with adjustments based on the patient and severity of the condition, to provide the greatest outcomes.⁽⁶⁾

Related work

Over the last few years, the majority of scientists have embraced and developed new various ML techniques that aim to predict cancer accurately and effectively.⁽⁷⁾ Many efforts have been spent on breast cancer diseases to develop and test ML techniques on different datasets.⁽⁸⁾ This section presents some of the research work related to this field with the type of data exploited, the predictive models used, and the performance measurement.

Uma Ojha et al.⁽⁹⁾ have compared various classifier algorithms on WPBC dataset. Their results demonstrate the classification algorithms; C5.0 and SVM have shown 81 % accuracy in classifying the recurrence of the disease.

Vikas Chaurasia et al.⁽¹⁰⁾ used three different data. The mining algorithms Naïve Bayes, RBF Network, and J48 were utilized to construct prediction models employing the Wisconsin dataset comprising 683 breast cancer cases. The models were assessed using a 10-fold cross-validation method to compare their performance. The findings revealed that Naïve Bayes emerged as the most accurate predictor with a precision of 97,36 %. Mohammed H. Tafish et al.⁽¹¹⁾ introduced a model aimed at addressing the challenge of assessing the risk level associated with the disease and optimizing healthcare practices, mitigation timelines, and costs. Their approach involved employing classification techniques such as SVM, ANN, and KNN on breast cancer data collected from hospitals in the Gaza Strip. Subsequent evaluation using these techniques yielded an accuracy rate of 77 %. Abdel-Zaher, A. M. et al.⁽¹²⁾ developed a CAD scheme for breast cancer detection utilizing a deep belief network unsupervised path followed by a supervised path employing backpropagation. The architecture integrates a back-propagation neural network with Liebenberg Marquardt learning function and weights initialized from the deep belief network path (DBNN). Their method, evaluated on the Wisconsin Breast Cancer Dataset (WBCD), achieved a classifier complexity resulting in an accuracy of 99,68 %.

METHOD

Data mining

Data mining is a step in the KDD process consisting of applying data analysis and discovery algorithms that, under acceptable computational efficiency limitations, produce a particular enumeration of patterns over the data. According to this definition, data mining is the step responsible for actual knowledge discovery.⁽¹³⁾

Software

Orange is an open-source data visualization, analysis, and machine learning toolkit. It offers a user-friendly interface for tasks such as data preprocessing, exploration, modeling, and evaluation. Orange provides a range of visual programming tools and widgets that allow users to create and execute data analysis workflows without needing to write code.⁽¹⁴⁾

Mammographic mass dataset

The dataset employed in this research comprises mammographic mass data and consists of a total of five attributes. Specifically, the breast cancer dataset referenced in the paper is derived from the larger Multi Cancer Dataset accessible on Kaggle. This subset of data comprises 10 000 images, evenly split between 5 000 malignant and 5 000 benign cases.⁽¹⁵⁾

Machine learning algorithms for prediction

The machine learning techniques utilized in this study incorporated a 10-fold cross-validation method, the dataset is divided into 10 subsets, with each subset serving as a test set once while the remaining data is used for training. This process is repeated 10 times to ensure a comprehensive assessment and mitigate any potential bias in the results. Various algorithms were utilized to assess classification accuracy (CA), arranged in ascending order from lowest to highest accuracy: Naïve Bayes with, Decision Tree, SVM, KNN, Logistic Regression, and Neural Network.^(16,17,18)

Here is more information about each algorithm used for analyses:

1. Naïve Bayes: Naïve Bayes is a probabilistic classification algorithm based on Bayes' theorem, with the assumption of independence between features. Despite its simplicity, it often performs well in practice, particularly for text classification tasks.^(19,20,21)
2. Decision Tree: A decision Tree is a popular algorithm used for classification and regression tasks. It recursively splits the dataset into subsets based on the most significant attribute at each node, resulting in a tree-like structure where leaves represent class labels or regression values.^(22,23,24)
3. Support Vector Machine (SVM): SVM is a supervised learning algorithm used for classification and regression tasks. It finds the hyperplane that best separates classes in the feature space, maximizing the margin between classes. SVMs are effective in high-dimensional spaces and when the number of features exceeds the number of samples.^(25,26,27)
4. K-Nearest Neighbors (KNN): KNN is a simple and intuitive algorithm used for classification and regression tasks. It classifies a data point based on the majority class among its k nearest neighbors in the feature space. KNN is non-parametric and lazy, meaning it does not require training and makes predictions based on stored instances.^(28,29,30)
5. Logistic Regression: Logistic Regression is a linear model used for binary classification tasks. It estimates the probability that a given input belongs to a particular class using the logistic function, which maps input features to the range [0, 1]. Despite its name, logistic regression is a classification algorithm, not a regression.^(31,32,33)
6. Neural Network: Neural Network is a flexible and powerful algorithm inspired by the biological neural networks of the human brain. It consists of interconnected layers of neurons, each performing a weighted sum of inputs followed by an activation function. Neural networks can learn complex patterns and relationships from data, making them suitable for a wide range of tasks.^(34,35,36,37)

RESULTS AND DISCUSSION

Table 1 provides an overview of the classification accuracy achieved by various algorithms for predicting breast cancer. In particular, the neural network algorithm performs well, recording an accuracy score of 0,978. Logistic regression follows closely behind, showing precision and a score of 0,965. The support vector machine (SVM) algorithm also performs well, achieving an accuracy of 0,967. In descending order, the K-Nearest Neighbors (KNN) algorithm achieves an accuracy of 0,943, while the decision tree model provides an accuracy of 0,884. Finally, the Naïve Bayes algorithm, although still popular among the tested methods, recorded the lowest accuracy with a score of 0,829. These findings show the different levels of success each algorithm can have in correctly predicting breast cancer and neural networks have emerged as the most advanced classifiers in this context.

Model	AUC	CA	F1	Prec	Recall	MCC	Spec
Neural Network	0,998	0,978	0,978	0,978	0,978	0,956	0,978
Logistic Regression	0,993	0,965	0,965	0,965	0,965	0,929	0,965
KNN	0,985	0,943	0,942	0,944	0,943	0,887	0,943
SVM	0,967	0,904	0,904	0,904	0,904	0,808	0,904
Tree	0,853	0,884	0,884	0,884	0,884	0,768	0,884
Naive Bayes	0,882	0,829	0,829	0,829	0,829	0,658	0,829

This section, elucidate the intended meaning of each abbreviation provided in the table.

1. AUC (Area Under the Curve): AUC measures the model’s ability to distinguish between positive and negative samples. A higher AUC indicates better performance.
2. CA (Classification Accuracy): CA represents the overall accuracy of the model in correctly classifying instances.
3. F1 Score: F1 score balances precision and recall, providing a single metric that considers both false positives and false negatives.
4. Precision (Prec): Precision measures the proportion of true positive predictions among all positive predictions.
5. Recall: Recall (also known as sensitivity or true positive rate) quantifies the proportion of actual positive instances correctly predicted by the model.

Confusion matrix

A confusion matrix is a table that is often used to evaluate the performance of a classification algorithm. It summarizes the performance of a classification model by presenting a breakdown of the number of correct and incorrect predictions made by the model on a set of data. The matrix consists of rows and columns where each row represents the actual class and each column represents the predicted class. The main diagonal of the matrix represents the instances that were correctly classified, while off-diagonal elements represent misclassifications.

		Predicted		
		breast_benign	breast_malignant	Σ
Actual	breast_benign	4879	121	5,000
	breast_malignant	97	4903	5,000
Σ		4,976	5,024	10,000

Table 2 displays the confusion matrix generated by the Neural network algorithm. It aligns with the results observed in table 1, indicating the superior performance of the Neural network. To elaborate, the model effectively identifies 4,903 malignant breast cases, with 121 cases erroneously classified as malignant. Furthermore, it accurately identifies 4,879 benign breast cases, with 97 cases misclassified as benign.

Confusion matrix values meaning and calculation:

- True Positives (TP): The number of correctly predicted malignant cases (4903).
- True Negatives (TN): The number of correctly predicted benign cases (4879).
- False Positives (FP): The number of benign cases incorrectly predicted as malignant (121).
- False Negatives (FN): The number of malignant cases incorrectly predicted as benign (97).

Now, let’s calculate some key metrics:

1. Accuracy:

$$\text{Accuracy} = (TP+TN) / \text{Total} = (4879+4903) / 10000 = 0,9782$$

2. Precision (Positive Predictive Value):

$$\text{Precision} = TP / (TP+FP) = 4903/(4903+121) = 0,9756$$

3. Recall (Sensitivity, True Positive Rate):

$$\text{Recall} = TP / (TP+FN) = 4903 / (4903+97) = 0,9806$$

4. F1 Score:

$$\text{F1 Score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) = 2 * (0,9756 * 0,9806) / (0,9756 + 0,9806) = 0,9781$$

5. Specificity (True Negative Rate):

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP}) = 4879 / (4879 + 121) = 0,9758$$

The manually calculated results indicate the validity of the results presented in the table, despite the presence of some minor differences.

Receiver Operating Characteristic (ROC)

Receiver Operating Characteristic (ROC) analysis is a method used to evaluate the performance of a binary classification model across different discrimination thresholds. It plots the true positive rate (sensitivity) against the false positive rate (1-specificity) for various threshold values, resulting in an ROC curve. The area under the ROC curve (AUC) provides a measure of the model's discriminative ability, with higher values indicating better performance. ROC analysis is widely used in clinical medicine and other fields as a fundamental evaluation tool for assessing diagnostic tests and predictive models. Figure 1 shows the ROC analysis for six different algorithms.

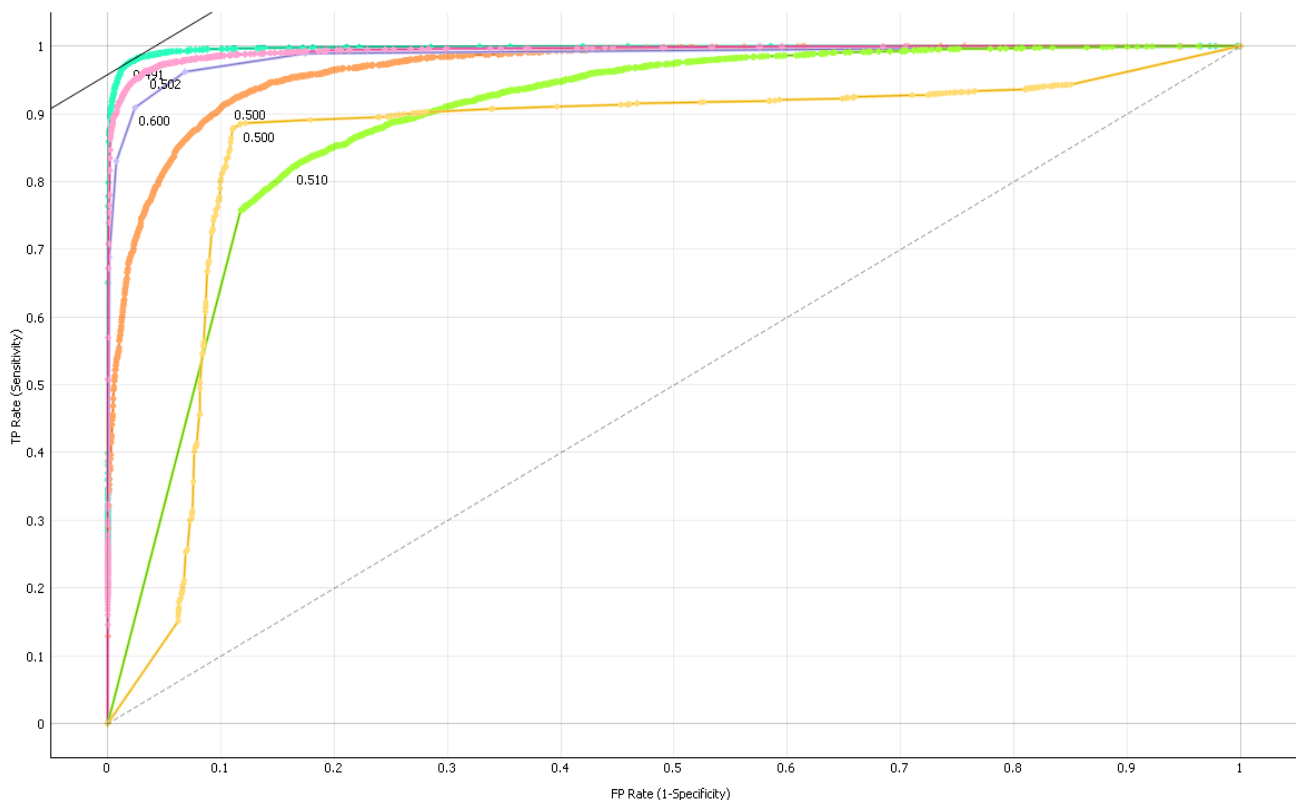


Figure 1. ROC analysis for six different algorithms

ROC analysis confirms the previous findings by indicating that the Neural network has the largest area under the curve. This reinforces the conclusion that the neural network model performs best in distinguishing between the classes, aligning with the results obtained earlier. Also, ROC analysis highlights a contrasting pattern with the Naïve Bayes model, indicating a smaller area under the curve compared to other models. This observation further supports the previous conclusion that the Naïve Bayes model may not perform as well in distinguishing between the classes, as evidenced by its lower discriminatory power.

CONCLUSION

In this study, we investigated six leading data mining techniques—Neural Networks, K-Nearest Neighbors (KNN), Naïve Bayes, Logistic Regression, Decision Trees, and Support Vector Machines (SVM)—to assess their efficacy in medical diagnosis and prognosis. The results of our comprehensive evaluation demonstrate that Neural Networks consistently achieved the highest accuracy, precision, and AUC metrics across the board, significantly outperforming traditional models. These findings confirm the significant role that Neural Networks and other advanced machine learning models can play in improving diagnostic precision and prognostic predictions. The superior performance of Neural Networks, in particular, highlights their potential for integration into clinical decision support systems, enabling healthcare professionals to make more informed, data-driven decisions. This not only accelerates the diagnosis process but also improves patient outcomes by facilitating early

intervention. While this study focuses on a comparative analysis of well-known classification techniques, future research should explore the integration of hybrid models and deep learning approaches to further enhance predictive capabilities. Additionally, real-world clinical trials and longitudinal studies are essential to validate the practical implementation of these models in diverse healthcare settings. Ultimately, the adoption of these data mining techniques can contribute to more accurate, timely, and personalized medical care, addressing one of the most critical challenges in modern medicine.

REFERENCES

1. Alzboon MS, Al-Batah M, Alqaraleh M, Abuashour A, Bader AF. A Comparative Study of Machine Learning Techniques for Early Prediction of Diabetes. In: 2023 IEEE 10th International Conference on Communications and Networking, ComNet 2023 - Proceedings. 2023. p. 1-12.
2. Alzboon MS, Al-Batah M, Alqaraleh M, Abuashour A, Bader AF. A Comparative Study of Machine Learning Techniques for Early Prediction of Prostate Cancer. In: 2023 IEEE 10th International Conference on Communications and Networking, ComNet 2023 - Proceedings. 2023. p. 1-12.
3. Al-shanableh N, Alzyoud M, Al-husban RY, Alshanableh NM, Al-Oun A, Al-Batah MS, et al. Advanced Ensemble Machine Learning Techniques for Optimizing Diabetes Mellitus Prognostication: A Detailed Examination of Hospital Data. *Data Metadata. 2024;3:363.*
4. Alqaraleh M, Abdel M. Advancing Medical Image Analysis : The Role of Adaptive Optimization Techniques in Enhancing COVID-19 Detection , Lung Infection , and Tumor Segmentation Avances en el análisis de imágenes médicas : el papel de las técnicas de optimización adaptativa para. *LatIA. 2024;2(74).*
5. Alawneh H, Hasasneh A. Survival Prediction of Children after Bone Marrow Transplant Using Machine Learning Algorithms. *Int Arab J Inf Technol. 2024;21(3):394-407.*
6. Alzboon MS, Alqaraleh M, Al-Batah MS. AI in the Sky: Developing Real-Time UAV Recognition Systems to Enhance Military Security. *Data Metadata. 2024;3(417).*
7. Ajay K, Sushil R, Tiwari A. Cancer Survival Analysis Using Machine Learning. *SSRN Electron J. 2019;*
8. Jafari-Marandi R, Davarzani S, Soltanpour Gharibdousti M, Smith BK. An optimum ANN-based breast cancer diagnosis: Bridging gaps between ANN learning and decision-making goals. *Appl Soft Comput J. 2018;72:108-20.*
9. Ojha U, Goel S. A study on prediction of breast cancer recurrence using data mining techniques. In: Proceedings of the 7th International Conference Confluence 2017 on Cloud Computing, Data Science and Engineering. 2017. p. 527-30.
10. Chaurasia V, Pal S, Tiwari BB. Prediction of benign and malignant breast cancer using data mining techniques. *J Algorithms Comput Technol. 2018;12(2):119-26.*
11. Tafish MH, El-Halees AM. Breast Cancer Severity Degree Predication Using Data Mining Techniques in the Gaza Strip. In: Proceedings - 2018 International Conference on Promising Electronic Technologies, ICPET 2018. 2018. p. 124-8.
12. Abdel-Zaher AM, Eldeib AM. Breast cancer classification using deep belief networks. *Expert Syst Appl [Internet]. 2016;46:139-44.* Available from: <https://doi.org/10.1016/j.eswa.2016.05.044>
13. Wahed MA, Alqaraleh M, Alzboon MS, Al-Batah MS. Application of Artificial Intelligence for Diagnosing Tumors in the Female Reproductive System: A Systematic Review. *Multidiscip. 2025;3:54.*
14. Ahmad A, Alzboon MS, Alqaraleh MK. Comparative Study of Classification Mechanisms of Machine Learning on Multiple Data Mining Tool Kits. *Am J Biomed Sci Res 2024 [Internet]. 2024;22(1):577-9.* Available from: www.biomedgrid.com
15. Alzboon MS, Al-Batah MS, Alqaraleh M, Abuashour A, Bader AFH. Early Diagnosis of Diabetes: A Comparison of Machine Learning Methods. *Int J online Biomed Eng. 2023;19(15):144-65.*

16. Al-Batah MS, Alzboon MS, Alzyoud M, Al-Shanableh N. Enhancing Image Cryptography Performance with Block Left Rotation Operations. *Appl Comput Intell Soft Comput.* 2024;2024(1):3641927.
17. Wahed MA, Alqaraleh M, Alzboon MS, Al-Batah MS. Evaluating AI and Machine Learning Models in Breast Cancer Detection: A Review of Convolutional Neural Networks (CNN) and Global Research Trends. *LatIA.* 2025;3:117.
18. Al-Batah M, Zaqaibeh B, Alomari SA, Alzboon MS. Gene Microarray Cancer classification using correlation based feature selection algorithm and rules classifiers. *Int J online Biomed Eng.* 2019;15(8):62-73.
19. Alqaraleh M, Alzboon MS, Al-Batah MS, Wahed MA, Abuashour A, Alsmadi FH. Harnessing Machine Learning for Quantifying Vesicoureteral Reflux: A Promising Approach for Objective Assessment. *Int J Online \& Biomed Eng.* 2024;20(11).
20. Al-Batah MS, Alzboon MS, Alazaidah R. Intelligent Heart Disease Prediction System with Applications in Jordanian Hospitals. *Int J Adv Comput Sci Appl.* 2023;14(9):508-17.
21. Alzboon MS. Internet of things between reality or a wishing-list: a survey. *Int J Eng \& Technol.* 2018;7(2):956-61.
22. Alzboon MS, Qawasmeh S, Alqaraleh M, Abuashour A, Bader AF, Al-Batah M. Machine Learning Classification Algorithms for Accurate Breast Cancer Diagnosis. In: 2023 3rd International Conference on Emerging Smart Technologies and Applications, eSmarTA 2023. 2023.
23. Alzboon M, Alomari SA, Al-Batah MS, Banikhalaf M. The characteristics of the green internet of things and big data in building safer, smarter, and sustainable cities. *Int J Eng \& Technol.* 2017;6(3):83-92.
24. Alzboon MS, Aljarrah E, Alqaraleh M, Alomari SA. Nodexl Tool for Social Network Analysis. Vol. 12, *Turkish Journal of Computer and Mathematics Education.* 2021.
25. Al Tal S, Al Salaimh S, Ali Alomari S, Alqaraleh M. The modern hosting computing systems for small and medium businesses. *Acad Entrep J.* 2019;25(4):1-7.
26. Alzboon MS, Al-Batah MS. Prostate Cancer Detection and Analysis using Advanced Machine Learning. *Int J Adv Comput Sci Appl.* 2023;14(8):388-96.
27. Al-Batah MS, Al-Eiadeh MR. An improved discreet Jaya optimisation algorithm with mutation operator and opposition-based learning to solve the 0-1 knapsack problem. *Int J Math Oper Res.* 2023;26(2):143-69.
28. Alzboon MS, Qawasmeh S, Alqaraleh M, Abuashour A, Bader AF, Al-Batah M. Pushing the Envelope: Investigating the Potential and Limitations of ChatGPT and Artificial Intelligence in Advancing Computer Science Research. In: 2023 3rd International Conference on Emerging Smart Technologies and Applications, eSmarTA 2023. 2023.
29. Al-Batah MS, Al-Eiadeh MR. An improved binary crow-JAYA optimisation system with various evolution operators, such as mutation for finding the max clique in the dense graph. *Int J Comput Sci Math.* 2024;19(4):327-38.
30. Alzboon MS, Bader AF, Abuashour A, Alqaraleh MK, Zaqaibeh B, Al-Batah M. The Two Sides of AI in Cybersecurity: Opportunities and Challenges. In: *Proceedings of 2023 2nd International Conference on Intelligent Computing and Next Generation Networks, ICNGN 2023.* 2023.
31. Al-Batah MS. Integrating the principal component analysis with partial decision tree in microarray gene data. *IJCSNS Int J Comput Sci Netw Secur.* 2019;19(3):24-29.
32. Alzboon M. Semantic Text Analysis on Social Networks and Data Processing: Review and Future Directions. *Inf Sci Lett.* 2022;11(5):1371-84.
33. Al-Batah MS. Testing the probability of heart disease using classification and regression tree model. *Annu*

Res Rev Biol. 2014;4(11):1713-25.

34. Alomari SA, Alqaraleh M, Aljarrah E, Alzboon MS. Toward achieving self-resource discovery in distributed systems based on distributed quadtree. J Theor Appl Inf Technol. 2020;98(20):3088-99.

35. Al-Batah MS. Ranked features selection with MSBRG algorithm and rules classifiers for cervical cancer. Int J Online Biomed Eng. 2019;15(12):4.

36. Alzboon MS. Survey on Patient Health Monitoring System Based on Internet of Things. Inf Sci Lett. 2022;11(4):1183-90.

37. Al-Batah MS. Modified recursive least squares algorithm to train the hybrid multilayered perceptron (HMLP) network. Appl Soft Comput. 2010;10(1):236-44.

FINANCING

This work is supported from Jadara University under grant number [Jadara-SR-Full2023], Zarqa University, and Applied Science Private University.

CONFLICT OF INTEREST

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Svitlana Stender, Olesia Lemishovska.

Data curation: Nataliia Roshko, Andrii Soloshchak.

Formal analysis: Svitlana Stender.

Research: Andrii Soloshchak.

Methodology: Olesia Lemishovska.

Project management: Olena Lagovska.

Resources: Nataliia Roshko, Olesia Lemishovska.

Software: Svitlana Stender.

Supervision: Olena Lagovska.

Validation: Svitlana Stender.

Display: Andrii Soloshchak.

Drafting - original draft: Olena Lagovska.

Writing - proofreading and editing: Svitlana Stender.