Development of a Web-Based Expert System for Early Tuberculosis Diagnosis Using K-Nearest Neighbors

Lukas Valentino [1], Vivi Nur Wijayaningrum [2], Candra Bella Vista [3]
* Politeknik Negeri Malang
lukasvalentino971@gmail.com [1], vivinurw@polinema.ac.id [2], bellavista@polinema.ac.id [3]

Article Info

Article history:

Received 2025-11-11
Revised 2026-02-19
Accepted 2026-04-10

Keyword:

Expert System,
KNN,
Machine Learning,
Tuberculosis Diagnosis.

ABSTRACT

Tuberculosis (TB) remains a major infectious disease that continues to challenge global public health, particularly in regions with limited access to early diagnostic services. Early and accurate detection of TB is essential to prevent further transmission and ensure effective treatment. This study aims to develop an intelligent expert system for early TB diagnosis based on the K-Nearest Neighbors (KNN) algorithm. To achieve this, the proposed system classifies patient conditions as TB or Non-TB based on clinical symptoms obtained from anonymized medical record data from a regional hospital in East Java, Indonesia. The dataset comprises eight symptom-based attributes, and several experiments were conducted to optimize parameters, including test-size ratio, number of neighbors (K), and K-Fold cross-validation. Results showed that the best model performance was achieved at $K = 5$ with a 10% test size, yielding an F1-score of 0.8571, precision of 1.00, and recall of 0.75. In addition, functional and non-functional testing confirmed that the developed web-based system operates reliably and securely. Overall, the proposed system provides an accurate and accessible diagnostic solution for early TB detection, especially in communities with limited healthcare infrastructure.

I. INTRODUCTION

Health is one of the most crucial aspects of human life, making the achievement and maintenance of good health a global priority. Among various health challenges, Tuberculosis (TB) remains one of the most serious infectious diseases worldwide [1]. TB is caused by Mycobacterium tuberculosis, a bacterium that primarily attacks the lungs but can also affect other organs. According to the World Health Organization (WHO), TB ranked as the second leading cause of death globally after COVID-19, with approximately 10 million people infected and 1.3 million deaths recorded in 2022 [2]. In many developing countries, including Indonesia, early diagnosis and treatment of TB are often hindered by limited healthcare facilities, a shortage of medical professionals, and low public awareness of TB symptoms, leading to delayed medical attention and higher mortality rates.

In response to these challenges, the use of computer-based decision support systems has gained increasing attention to support early disease detection. One promising approach is the development of an expert system, also known as a knowledge-based system, which is designed to replicate human expert reasoning in solving specific problems. This system operates by utilizing knowledge and analytical methods that have been predefined by experts according to their area of expertise, allowing it to simulate expert-level decision-making within a specific domain [3]. In the medical domain, expert systems can analyze patient symptoms and provide preliminary diagnostic suggestions, helping healthcare professionals and the general public, especially those in remote areas with limited access to medical facilities, to detect diseases earlier and seek appropriate treatment.

Several studies have investigated the application of machine learning algorithms in medical diagnosis, particularly for disease classification and prediction tasks. For instance, Wijayaningrum optimized Multi-Layer Perceptron (MLP) parameters for early-stage diabetes prediction and demonstrated that parameter tuning using metaheuristic algorithms could significantly improve diagnostic accuracy

[4]. Assegie proposed a heart disease prediction model using the K-Nearest Neighbor (KNN) algorithm and demonstrated that KNN achieved an accuracy of 91.99%, indicating its reliability in medical classification compared to other algorithms such as Naïve Bayes, Decision Tree, and Random Forest [5]. Lestari also compared KNN with Support Vector Machine (SVM) for cardiovascular disease prediction using data from the UCI Machine Learning Repository. Their results showed that KNN provided competitive accuracy and stable performance, highlighting its ability to classify medical data with minimal parameter tuning [6].

Subsequent research continued to validate KNN's robustness. Maheswari analyzed breast cancer image classification using both KNN and Naïve Bayes classifiers and found that KNN achieved higher accuracy and F1-scores, confirming its strength in handling feature-rich medical datasets [7]. Similarly, Sravani applied KNN for cardiovascular disease detection and obtained an accuracy of 90%, outperforming Naïve Bayes and demonstrating better consistency across precision, recall, and F1 metrics [8]. In the same year, Madani also enhanced the traditional KNN algorithm by introducing a Weighted KNN (WKNN) model with Gaussian kernel weighting, achieving 91.8% accuracy in heart disease prediction and demonstrating the potential of KNN variants to improve diagnostic precision [9]. Pratama conducted a comparative analysis of KNN performance in medical data classification and confirmed that the algorithm achieved stable and competitive accuracy across multiple evaluation metrics [10]. Likewise, Ghaisani implemented the KNN Regressor method to predict Toyota used car prices and obtained high prediction accuracy with low error values, confirming KNN's adaptability and effectiveness in regression-based prediction tasks [11].

Further, Nasution developed a system combining KNN and Convolutional Neural Network (CNN) algorithms to identify bronchitis disease in toddlers using thorax X-ray images. Their study revealed that the KNN classifier achieved 92% precision and 99% recall when using Manhattan distance metrics, confirming its effectiveness for early medical diagnosis [12]. Lewandowicz also compared KNN, Naïve Bayes, and SVM algorithms for heart disease classification and concluded that KNN outperformed Naïve Bayes in terms of accuracy, precision, recall, and F1-score, highlighting its robustness and adaptability for health data prediction [13]. Furthermore, Munira applied the KNN algorithm in a web-based expert system for diagnosing exanthema virus infections. The study showed that KNN was effective in distinguishing diseases with similar symptoms and provided accurate diagnostic results comparable to human experts [14]. Collectively, these studies demonstrate that KNN consistently performs well across diverse medical datasets and diagnostic contexts [15], [16], [17]. Its non-parametric nature, simplicity, and strong generalization capability make it a reliable choice for disease classification models.

Despite these advancements, few studies have focused on implementing KNN-based expert systems for TB diagnosis, particularly for self-diagnosis applications that empower users to identify TB symptoms independently before consulting medical experts. Most existing studies are limited to general medical prediction models and lack integration with user-friendly platforms suitable for community-level utilization. Hence, there remains a research gap in developing an intelligent and accessible diagnostic system specifically tailored to TB detection using clinical symptom data.

Unlike previous studies that primarily focused on general disease prediction or standalone classification models, this research integrates a KNN-based classifier into a web-based expert system specifically designed for early TB self-diagnosis using clinical symptom data. Therefore, this research proposes the development of an Expert System for Tuberculosis Diagnosis using the KNN algorithm. The system is designed to classify patients as TB or non-TB based on eight clinical symptoms, namely productive cough, fever, night sweats, weight loss, loss of appetite, malaise, shortness of breath, and chest pain. The KNN algorithm was selected due to its non-parametric nature, robustness in handling small-to-moderate datasets, minimal training complexity, and competitive performance reported in various medical classification studies. The objectives of this research are to develop a web-based expert system capable of performing early TB detection based on patient symptoms and evaluate the classification performance of the KNN algorithm in diagnosing TB using anonymized medical record data from a regional hospital in East Java, Indonesia. This system is expected to serve as a supportive tool for early TB detection and contribute to improving public health outcomes, especially in areas with limited healthcare resources.

II. METHODS

This study followed the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework, which is widely used for data mining and machine learning research. The framework provides a systematic, iterative process consisting of six key stages, as shown in Figure 1. This approach was selected as it ensures a comprehensive workflow from problem identification to model validation and system implementation.

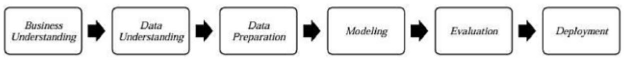

Figure 1. Research framework based on CRISP-DM

A. Data Collection and Description

The dataset used in this study consisted of 180 anonymized medical records obtained from a regional hospital in East Java, Indonesia. This sample size was considered sufficient for preliminary model evaluation and validation within a controlled experimental setting. The dataset consisted of 90 TB cases and 90 Non-TB cases, ensuring representation of both classes for classification analysis. Each record contained eight symptom-based attributes and one target label representing the physician's diagnosis of either Tuberculosis

(TB) or Non-TB. The symptoms included productive cough, fever, night sweats, weight loss, loss of appetite, malaise, shortness of breath, and chest pain. All patient identifiers were removed to maintain confidentiality, and the dataset was handled following ethical research principles. A brief description of each feature is presented in Table 1.

TABLE I
DESCRIPTION OF DATASET ATTRIBUTES

Parameter	Description
Productive Cough	Persistent cough producing phlegm for more than two weeks
Fever	Body temperature between 38.5–39.5°C without clear cause
Night Sweats	Excessive sweating at night without physical activity
Weight Loss	Unexplained weight loss over several weeks
Loss of Appetite	Reduced appetite leading to lower nutrient intake
Malaise	Feeling of fatigue or weakness for more than two weeks
Shortness of Breath	Difficulty in breathing during daily activities
Chest Pain	Pain or discomfort in the chest when breathing or coughing

B. Data Preparation

Prior to model training, several preprocessing steps were performed to ensure the dataset was clean, consistent, and suitable for machine learning analysis. First, data cleaning was performed to eliminate duplicate entries and handle any missing values. Next, categorical symptom values expressed as “yes” or “no” were encoded into numerical form to allow algorithmic processing. Each of the eight clinical symptoms served as an input feature, while the diagnostic label (TB or Non-TB) acted as the output variable. To prevent bias caused by differing feature scales, normalization was applied so that all attributes contributed equally to distance calculations during classification. Finally, the data were divided into training and testing subsets using random sampling, allowing the model to learn from one portion of the data and be validated on another.

C. Modeling Using K-Nearest Neighbors

The predictive model in this study was developed using the *K*-Nearest Neighbors (KNN) algorithm, a simple yet effective method for classification tasks. KNN determines the class of a new instance based on its similarity to existing labelled instances in the dataset. The algorithm identifies a specific number of neighboring samples, denoted as *K*, that are closest to the input case and assigns the class most frequently represented among them. The closeness between samples was measured using a standard distance metric, ensuring that instances with higher similarity contributed more strongly to the classification decision.

KNN was chosen due to its proven reliability in various medical diagnosis studies, especially when the dataset size is moderate and the feature relationships are not linearly separable. Its advantages include robustness to small data fluctuations and ease of implementation. However, the method can be sensitive to outliers and requires properly normalized data for optimal performance. In this research, several values of *K* were later tested during the validation stage to determine the most accurate configuration for TB diagnosis.

D. Model Evaluation

To assess the effectiveness of the classification model, three commonly used performance metrics, Precision, Recall, and F1-Score, were applied. Precision measures the accuracy of the model’s positive predictions, indicating how many of the patients classified as TB were actually diagnosed with TB. Recall reflects the model’s ability to identify all true TB cases within the dataset, minimizing missed detections. The F1-Score provides a harmonic mean between Precision and Recall, providing a single value that summarizes the overall quality of the classification results.

In medical applications, maintaining a balance between these metrics is essential, as both false positives and false negatives can lead to significant diagnostic consequences. Therefore, the F1-Score was selected as the primary indicator of model performance, with higher values representing more accurate and dependable classifications.

E. System Implementation

Following model development and evaluation, the final system was implemented as a web-based expert system to facilitate user interaction and diagnosis accessibility. The machine-learning model was developed using the Python programming language, while the Flask framework was utilized to provide an application programming interface (API) connecting the trained model to the system’s front end. The user interface was built using the Laravel framework to ensure scalability and ease of maintenance.

Two main user roles were defined, namely the Administrator, responsible for managing user accounts, symptom data, and system configurations; and the Patient/User, who can input experienced symptoms and receive immediate diagnostic feedback. Once the user submits symptom data, the system processes the input through the trained KNN model and displays the classification result, either “TB” or “Non-TB”.

III. RESULT AND DISCUSSION

This section presents the experimental results and performance evaluation of the proposed Expert System for Tuberculosis Diagnosis using the KNN algorithm. Several tests were conducted to determine the optimal configuration of model parameters and to evaluate both the functionality and reliability of the implemented system. The evaluation includes parameter optimization tests (Train-Test Split, *K*

Value, and *K*-Fold Cross Validation), training–testing performance analysis, and system-level testing (functional and non-functional).

A. Train–Test Split Analysis

The first experiment investigated the effect of varying the proportion of test data on the model's performance. The test size ratio was varied between 90:10 and 10:90 [18], and the resulting F1-score, precision, and recall values were compared to identify the optimal configuration. The comparison of F1-score at various test size ratios is presented in Figure 2.

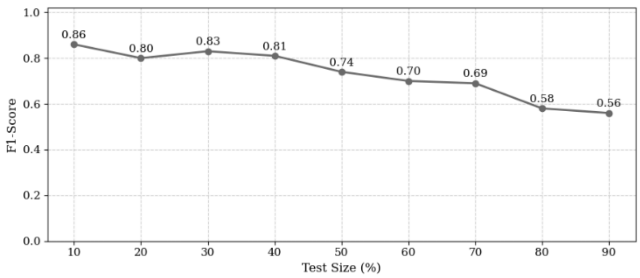

Figure 2. Comparison of F1-score at various test size ratios

The results showed that the highest F1-score (0.8571 or 86%) was achieved when the test size was 10%. This configuration produced a perfect precision value of 1.00, indicating that all instances predicted as TB were correctly classified, with no false positives. However, the recall value for the TB class was 0.75, meaning that two out of eight TB cases were not detected by the model. In medical diagnosis, a lower recall indicates that some TB cases may remain undetected, which emphasizes the need for further dataset expansion and model refinement to improve sensitivity. Increasing the test data proportion led to a decline in model performance, as the reduced training data limited the algorithm's ability to generalize symptom patterns. These findings indicate that a smaller test proportion allows the model to learn more effectively, leading to better classification results. Therefore, a 10% test size was selected as the optimal data split for subsequent testing.

B. K Value Optimization

The second experiment aimed to determine the optimal number of nearest neighbors (*K*) that yielded the highest F1-score. The test was conducted using a 10% test size and an automatic early-stopping mechanism to halt the process once no significant improvement in F1-score was observed across several iterations. The relationship between the *K* value and the resulting F1-score of the model is illustrated in Figure 3.

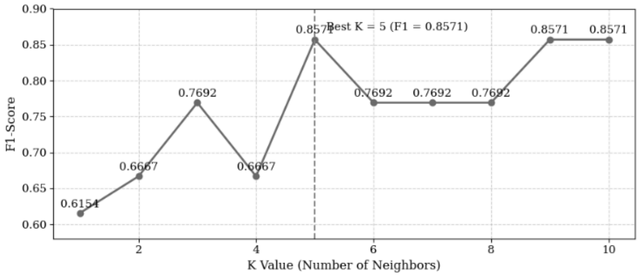

Figure 3. F1-Score comparison across different K values

The results revealed that the maximum F1-score of 0.8571 (86%) was achieved at *K* = 5, *K* = 9, and *K* = 10. Although these values produced equivalent results, *K* = 5 was selected as the optimal configuration because it provided a balance between accuracy and computational efficiency. Using a smaller *K* value helps the model remain sensitive to subtle variations in the data, while excessively large *K* values may smooth out the classification boundary and lead to overgeneralization. This outcome suggests that the model performs consistently within the *K* range of 5–10, demonstrating robustness against variations in neighborhood size.

C. K-Fold Cross Validation

The third experiment used *K*-Fold Cross Validation to evaluate the model's stability across multiple data partitions. In this test, the number of folds varied from 5 to 10 [18] while maintaining *K* = 5. The results confirmed that the model performed consistently across all configurations, with only minor variations in average F1-score. The variation of average F1-scores across different *K*-Fold configurations is shown in Figure 4.

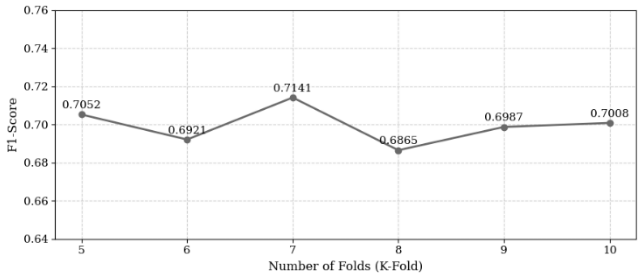

Figure 4. Average F1-score comparison across different fold values

The best cross-validation performance was obtained with 7 folds, yielding an average F1-score of 0.7141 (71.41%). The 5-fold and 10-fold configurations produced slightly lower averages of 0.7052 and 0.7008, respectively. The F1-score range across folds (0.625–0.7857 for 5-fold and 0.5–0.8889 for 10-fold) indicated reasonable variability and stable predictive capability. These results demonstrate that the model maintained consistent performance regardless of data division, confirming the general reliability of the KNN classifier for TB diagnosis under different validation settings.

D. Analysis of Training and Testing Performance

Training and testing experiments were conducted to further analyze the model's ability to generalize beyond the training data. Both datasets were evaluated using various test-size configurations to assess differences in performance stability. The comparison between training and testing F1-scores for various test size ratios is depicted in Figure 5.

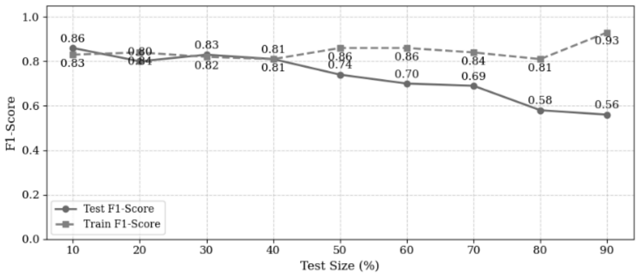

Figure 5. Comparison of training and testing F1-score across test sizes

As shown in Figure 5, the model achieved its best results at a test size of 10%, where the testing F1-score reached 0.8571, slightly higher than the training F1-score of 0.8323. This small gap indicates good generalization capability, suggesting that the model can maintain high predictive accuracy even when exposed to unseen data. For the TB class (Class 1), the model achieved perfect precision (1.00), meaning all positive TB predictions were correct. Conversely, the Non-TB class (Class 0) achieved perfect recall (1.00), indicating no Non-TB cases were mistakenly classified as TB. However, the recall for the TB class remained at 0.75, implying that a few TB cases were missed by the model. This analysis confirms that the KNN algorithm effectively recognizes most TB symptom patterns but could benefit from additional data enrichment or feature weighting to improve sensitivity toward less common cases.

E. System Testing and Validation

System testing was conducted to verify that the developed expert system operated in accordance with its functional and non-functional requirements. The testing process aimed to ensure that the system not only performed accurate diagnostic computations using the KNN algorithm but also met the required standards of usability, reliability, and performance.

From the functional perspective, the main features of the system, including symptom data input, processing through the KNN classification model, and generation of diagnostic results, were tested comprehensively using real symptom combinations. The results confirmed that the system successfully produced accurate diagnostic outcomes consistent with the expected logic. Additional components such as user authentication, symptom management, and role-based access control also functioned without error. These results indicate that the system's functional components have been effectively designed and implemented, demonstrating stable performance across different usage scenarios.

In terms of non-functional validation, the system's performance was assessed based on its availability, security, and compatibility. The expert system demonstrated high availability, operating smoothly in an online environment with stable response times as long as an internet connection was present. Security testing confirmed that user credentials and sensitive data were securely managed through encryption, while access control mechanisms successfully restricted unauthorized usage. Compatibility testing further verified that the system ran seamlessly on various popular web browsers such as Google Chrome, Mozilla Firefox, and Microsoft Edge, with no interface distortion or functionality issues.

Overall, the combined results from both functional and non-functional testing indicate that the proposed expert system is stable, secure, and functionally reliable. It is therefore suitable for broader use in assisting early self-diagnosis of Tuberculosis, particularly in regions with limited access to medical facilities.

F. Accuracy Validation with Expert Diagnosis

To validate the reliability of the system in real-world scenarios, the model's diagnostic results were compared with expert assessments from two physicians. Using 15 test samples, the system achieved a 100% match with the physicians' primary diagnoses of Pulmonary TB. This result reflects agreement within the limited validation subset and does not contradict the recall variability observed during model evaluation.

However, in three cases, the physicians provided additional diagnoses indicating comorbid conditions such as Pneumonia, Malnutrition, COPD, and Rheumatism, which were not detected by the system. This limitation highlights that the current model can accurately identify primary TB infections but requires further enhancement to recognize overlapping or secondary diseases. Overall, this evaluation demonstrates that the KNN-based expert system achieved strong performance in detecting Pulmonary TB, meeting the primary objective of this research: to support early and accurate self-diagnosis, particularly in regions with limited access to medical services.

IV. CONCLUSION

This study successfully developed and evaluated an Expert System for Tuberculosis Diagnosis based on the KNN algorithm. The system was designed to assist users in performing early self-diagnosis through symptom-based classification. Experimental results demonstrated that the model achieved a high level of accuracy, with an F1-score of 0.8571 under optimal parameters (test size of 10% and $K = 5$) and a stable performance across different K-Fold Cross Validation configurations. The validation against medical expert diagnoses showed a 100% match for the primary disease, confirming the model's effectiveness in recognizing Pulmonary TB cases. Furthermore, system testing verified that both functional and non-functional aspects met the expected requirements, indicating that the proposed system is stable, secure, and ready for initial deployment as a decision-support tool for early TB detection, especially in regions with limited healthcare access.

Future research should focus on expanding and enriching the dataset by incorporating a larger and more diverse set of medical records, including additional features such as cough duration, sputum characteristics, and chest X-ray indicators. This improvement is essential to enhance the model's sensitivity and allow it to identify comorbid conditions, which were not identified in the current version of the system. It is also recommended to compare the KNN algorithm with other machine learning methods such as Random Forest, Decision Tree, or Support Vector Machine to determine which algorithm provides the best balance between accuracy and computational efficiency. Moreover, integrating the expert system into a mobile application platform could significantly increase accessibility and usability for communities in remote areas, thereby strengthening the system's practical contribution to public health monitoring and early disease prevention.

REFERENCES

[1] Y. Chen *et al.*, "Global burden of HIV-negative multidrug- and extensively drug-resistant tuberculosis based on Global Burden of Disease Study 2021," *Science in One Health*, vol. 3, p. 100072, Jan. 2024, doi: 10.1016/j.soh.2024.100072.

[2] O. B. Kwon *et al.*, "Nutrition Status and Comorbidities Are Important Factors Associated With Mortality During Anti-Tuberculosis Treatment," *J Korean Med Sci*, vol. 40, no. 17, p. e73, 2025, doi: 10.3346/jkms.2025.40.e73.

[3] Z. Indra, "Development of an Expert System To Diagnose Tb in Pandemi Time Using the Forward Chaining Method in Puskesmas Medan, Johor," *Karismatika*, vol. 2, no. 1, pp. 40–46, 2022.

[4] V. N. Wijayaningrum, T. H. Saragih, and N. N. Putriwijaya, "Optimal multi-layer perceptron parameters for early stage diabetes risk prediction," in *IOP Conference Series: Materials Science and Engineering*, 2021, p. 012070. doi: 10.1088/1757-899x/1073/1/012070.

[5] T. A. Assegie, "Heart disease prediction model with k-nearest neighbor algorithm," *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 10, no. 3, pp. 225–230, Dec. 2021, doi: 10.11591/ijict.v10i3.pp225-230.

[6] W. Lestari and S. Sumarlinda, "Implementation of K-Nearest Neighbor (KNN) And Support Vector Machine (SVM) For Clasification Cardiovascular Disease," *International Journal of Multi Science*, vol. 2, no. 10, pp. 30–36, 2022, [Online]. Available: https://archive.ics.uci.edu/ml/datasets/heart+disease.

[7] B. U. Maheswari, T. Guhan, C. F. Britto, A. Sheeba, M. P. Rajakumar, and K. Pratyush, "Performance analysis of classifying the breast cancer images using KNN and naive bayes classifier," in *AIP Conference Proceedings*, AIP Publishing LLC, 2023, p. 020012. doi: https://doi.org/10.1063/5.0164139.

[8] S. Sravani and P. R. Karthikeyan, "Detection of cardiovascular disease using KNN in comparison with naive bayes to measure precision, recall and f-score," in *AIP Conference Proceedings*, AIP Publishing LLC, 2023, p. 030002. doi: https://doi.org/10.1063/5.0177014.

[9] F. Madani, K. Kusworo, and F. Farikhin, "Heart Disease Prediction Using Optimized Weighted K-Nearest Neighbor (WKNN)," *Jurnal Penelitian Pendidikan IPA*, vol. 10, no. 11, pp. 8847–8854, Nov. 2024, doi: 10.29303/jppipa.v10i11.9257.

[10] Y. D. Pratama and A. Salam, "Comparison of Data Normalization Techniques on KNN Classification Performance for Pima Indians Diabetes Dataset," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 3, pp. 693–706, 2025, [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[11] M. S. Ghaisani and A. Baita, "Implementation of the K-Nearest Neighbors (KNN) Regressor Method to Predict Toyota Used Car Prices," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 1, pp. 194–201, 2025, [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[12] M. F. Nasution, W. Wanayumini, and R. Roesnelly, "Classification of K-Nearest Neighbor (K-NN) and Convolutional Neural Network (CNN) for the Identification of Bronchitis Disease in Toddlers Using GLCM Feature Extraction Based on Thorax X-Ray Images," in *Proceeding of International Conference on Education, Society And Humanity*, Mar. 2024, pp. 1637–1648.

[13] B. Lewandowicz and K. Kisiała, "Comparison of Support Vector Machine, Naive Bayes, and K-Nearest Neighbors Algorithms for Classifying Heart Disease," in *International Conference on Information and Software Technologies*, Cham: Springer Nature Switzerland, 2024, pp. 274–285. doi: https://doi.org/10.1007/978-3-031-48981-5_22.

[14] Z. Munira, H. Husaini, and Z. Khalid, "Expert System for Diagnosing Exanthema Virus Using Web-Based KNN (K-Nearest Neighbor) Method," *InnoComp : Journal of Informatics and Computing*, vol. 1, no. 1, pp. 17–24, 2025, [Online]. Available: https://ejournal.sagita.or.id/index.php/InnoComp

[15] I. Khasanah, "Enhancing Alzheimer's Disease Diagnosis with K-NN: A Study on Pre-processed MRI Data," *International Journal of Artificial Intelligence in Medical Issues*, vol. 2, no. 1, pp. 49–60, May 2024, doi: 10.56705/ijaimi.v2i1.150.

[16] A. A. Permana and A. Arsanah, "Enhancing Early Diagnosis of Heart Disease: A Comparative Study of K-NN and Naive Bayes Classifiers Using the UCI Heart Disease Dataset," *Journal of Intelligent Computing & Health Informatics*, vol. 5, no. 1, pp. 35–42, Apr. 2024, doi: 10.26714/jichi.v5i1.11251.

[17] Ari Peryanto, D. Susanto, and B. H. Jihad, "Classification of Skin Disease Images Using K-Nearest Neighbour (KNN)," *Journal of Advanced Health Informatics Research*, vol. 2, no. 3, pp. 168–174, Feb. 2025, doi: 10.59247/jahir.v2i3.300.

[18] V. N. Wijayaningrum, A. P. Kirana, I. K. Putri, and T. O. Satrio, "Prediction of Student Academic Performance in Practicum Courses Based on Activity Logs and Student Background," *Proceedings - IEIT 2022: 2022 International Conference on Electrical and Information Technology*, pp. 366–371, 2022, doi: 10.1109/IEIT56384.2022.9967888.