Pujala Sivaleela, 2026, 14:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

# An Intelligent Healthcare Prediction System Using Ensemble Machine Learning and Generative AI

**Pujala Sivaleela[1], M.Prasanna Kumar[2]**

[1]Student, [2]Assistant Professor, Department of CSE,AI,ML, MAM Women's Engineering College Kesanupalli Narasaraopet

**Abstract- Today, because to advancements in both technology and healthcare, people may have access to affordable, personalised health assistance right where they are. This research presents a method that may help individuals with their health by giving them detailed information on their health and being able to forecast when they may become sick. In order to foretell potential health problems, users are asked to provide three main symptoms together with basic biographical details such as gender and age. Decision trees, random forests, naive bayes, logistic regression, support vector machines (SVMs), K-nearest neighbours, and XGBoost are the seven machine learning models that the system uses to provide a final prediction using a voting ensemble model. In order to provide thorough descriptions of the expected illness, including its origins, symptoms, potential treatments, and remedies, the system integrates generative AI with predictive analytics. The generative AI model takes into account the user's age and gender to provide information that is personalised to their requirements. The reliability and effectiveness of the system may be shown by evaluation metrics including F1-score, accuracy, precision, and recall. People are better able to make informed choices about their health because to our user-friendly platform that makes health information freely accessible.**

**Keywords: illness prediction, voting ensemble, generative AI, machine learning, and predictive models.**

## I. INTRODUCTION

Modern advances in generative artificial intelligence (GAI) and machine learning (ML) are reshaping health management and comprehension. Obtaining fast, affordable, and personalised health information is still challenging, despite recent developments. A health assistant that provides detailed information on possible diagnosis and lets users predict illnesses based on symptoms is presented in this research. Through the integration of GAI and ML, the technology offers a tailored, frictionless experience that places an emphasis on ease of use. In order to provide predictions that are strong and applicable to various populations, future updates will try to incorporate a wider range of demographic and geographical variables, even if the present dataset only covers 42 illnesses and 128 symptoms.

Inputs from the user initiate the prediction process: Please include your age, gender, and three symptoms. The system uses seven machine learning algorithms—XGBoost, Decision Tree, Random Forest, Naive Bayes, SVM, and Logistic Regression—to handle these inputs. To provide a thorough knowledge base, each model is trained independently on a dataset that includes 42 illnesses and 128 symptoms. The voting en semble approach achieves an ideal balance between the strengths of several models by combining their forecasts. This results in a final, consensus-based prediction that is more reliable and less biassed.

Future work will investigate how to make the ensemble technique scalable, especially in contexts with limited resources, so that mobile and edge devices may be deployed efficiently. The GAI module improves the user experience by giving a detailed, personalised overview of the situation after a potential illness has been identified. By using Google's GAI API, the system offers comprehensive health information, including descriptions, causes, symptoms, treatments, and treatment choices. An essential part is personalisation, which makes the shared health information even more relevant by adapting each answer to the user's gender and age.

## II. RELATED WORKS

The rapid development of healthcare technology, especially in areas such as artificial intelligence and machine learning, is contributing to the finding of new answers to the growing number of problems caused by health difficulties. The implications of these advancements for precise diagnosis, illness prediction, and personalised healthcare solutions are substantial. This review paper integrates earlier research from several domains, including rule-based systems, GAI, and semantic web based AI approaches, to emphasise the ways these technologies affect illness diagnosis, treatment, and patient care.

Recently, there have been notable advancements in illness prediction made possible by artificial intelligence methodology sourced from the semantic web. One such approach is the integration of disease ontology with rule-based reasoning. By synthesising ontological explicit information into the Semantic Web Rule Language (SWRL), the innovative approach developed by Dongre et al. ML toGAI [1] uses ChatGPT and ontology to provide accurate forecasts with explanations that patients can understand.

To bridge the gap between AI's predictive powers and the healthcare sector's desire for interpretability, this approach emphasises the need to augment ML with semantic technologies. Particularly effective in medical AI applications employing rule-based reasoning is the combination of domain-specific knowledge with traditional ML methods. For a very accurate deduction, Zolhavarieh et al. [2] used SWRL to categorise TB based on clinical signs. There is a lot of overlap in symptoms across illnesses, hence language-based techniques (rule-based) are strong in disease diagnosis. In a recent study, Chandra et al. [3] employed disease ontology with RDF and SWRL to categorise vector-borne diseases.

By enhancing image quality, synthesising high-quality photos, and allowing enhanced diagnostic procedures, GAI—a subset of AI—has contributed substantial breakthroughs to the field of scientific imaging. Some popular generative models used in healthcare include variational autoencoders (VAEs) and generative adversarial networks (GANs) [4]. Research by Manubolu and Peeriga [5] investigated GANs' potential for scientific image synthesis, drawing attention to the ways in which they may improve evaluation and reduce noise in imaging modalities like as CT and MRI.

By allowing early diagnosis and action, this capacity not only aids radiologists in analysis but also optimises patient outcomes. Certain image-processing applications of GAI, including picture segmentation and artefact reduction, have enabled oncology to rely on precise lesion detection [6]. When training ML models, generative models have made it much simpler to create synthetic pictures, which is especially helpful when there is a lack of real-world data [7].

Using GANs to generate synthetic data has been very helpful in diagnostic imaging, where high-quality data is needed to improve the accuracy of model training and forecasts. Combining rule-based techniques with ontology is an effective method for finding and classifying diseases with complex or overlapping symptoms. El Mas sari et al. [8] used ontology models and SWRL rules to improve classification accuracy and develop an explainable AI model for diabetes forecasting.

Clear SWRL criteria made the categorisation procedure easy to grasp, which improved communication and understanding between doctors and their patients. Several investigations have investigated the possibility of integrating machine learning with ontology-driven rule-based systems to enhance diagnostic capabilities. Through the use of automated categorisation using clinical indicators, Bensalah et al. [9] shown how SWRL and ontology-based inference might enhance the identification of bone tumours. By including ontology into their model, they demonstrated how this kind of system may connect clinical diagnostics with outcomes that patients can comprehend, allowing doctors to better communicate complicated information.

Improving Patient Outcomes with AI That Is Easy to Understand An increasingly pressing requirement in healthcare is the use of explainable AI to provide transparency and accountability in medical decision-making. The work of Ribeiro et al. [10], who developed Local Interpretable Modelagnostic Explanations (LIME), has seen extensive use in providing context for the predictions made by complex AI models. Despite LIME's effectiveness in highlighting the importance of characteristics, research has shown that interpretability at the patient level is still somewhat restricted.

By using ChatGPT for patient-level explanations, MLtoGAI is able to circumvent this limitation and provide transparent, personalised health advice [1]. In similar experiments, ChatGPT has shown promise in helping patients better understand medical conditions and recommended treatments by translating model predictions into practical, understandable insights that humans can use [11]. The interpretability of ChatGPT, when integrated with SWRL-based rule systems, is a huge step forward in personalised healthcare solutions that focus on the patient.

Comparative studies in AI-driven diagnostics have shown that rule-based and ML-based approaches each have their own set of benefits and drawbacks, which vary with the illness situation. For example, although traditional ML approaches have excelled in classifying massive datasets, rule-based systems really shine when it comes to issues of interpretability and transparency. In disciplines like infectious diseases and cancer, where clinical judgement is mostly reliant on symptom-based inference, rule based techniques are preferable, according to both Devi et al. [12] and El Massari et al. [8]. As far as the performance of ML models is concerned, hybrid models that combine ML with rule-based systems have shown promising results in terms of achieving high prediction accuracy while maintaining interpretability. Demonstrating the efficacy of combining algorithmic reasoning with ontological information, hybrid systems have been developed for vector-borne illnesses (Sandh et al., [13]) and dengue (Devi et al., [12]).

Semantic web technologies, generative AI, and rule-based AI are often used together in medical applications, according to the literature. By achieving a happy medium between diagnostic precision and openness, these systems hope to fulfil healthcare's dual objectives of effectiveness and interpretability. Hybrid AI models that integrate GAI's flexibility and predictive power with rule-based systems' interpretability should be the focus of future research, especially for illnesses with complex symptomatology and a high desire for personalised patient communication.

## III. METHODOLOGY

Applying several machine learning models to symptom-based illness forecasting is the objective of this study. Also, by giving detailed information about the illnesses that generative AI predicts, it improves the user experience. The next techniques address data analysis, model training, and interaction with a GAI API in depth. They also cover the ensemble approach. Part A: Priming and Data Loading

One goal column named "prognosis" organises the dataset, and it comprises 42 distinct sickness classifications. There are 128 columns that represent various symptoms. The illnesses are listed in a categorised fashion in the "prognosis" column. There is a binary value for each symptom; 1 means the symptom is present, and 0 means it is not. For training purposes, the dataset is kept clean by removing any unnecessary columns, such as

Unnamed: 133. Label encoding is used to the "prognosis" column to transform the illness names into numerical values that may be used by machine learning algorithms. Finally, the train test split command divides the dataset in half, with half going to the training set and the other half to the testing set. This allows for an objective evaluation and helps the models to efficiently generalise to new, untested data.

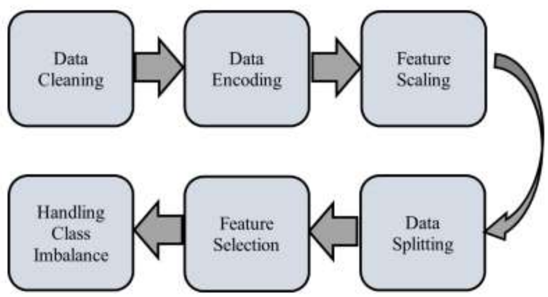


Fig. 1. Data Preprocessing Workflow

**Section B: Training and Model Selection**

Seven models were chosen to capture different types of data prediction patterns:

- A decision tree is a feature-based tree model that produces feature-based splits, which improves interpretability. To improve the accuracy of forecasts, a Random Forest combines many decision trees and averages their results. The Naive Bayes model is a good example of a probabilistic model used for tasks involving the classification of discrete features.
- **Logistic Regression:** For classification problems, it is successful since it is a linear model that predicts prob abilities for binary outcomes using a sigmoid function.

One effective method for high-dimensional domains is the Support Vector Machine (SVM), which maximises the decision boundary between classes. One distance-based model that uses the characteristics' nearest neighbours to generate predictions is K-Nearest Neighbours (KNN). One gradient boosting approach that builds trees sequentially to enhance accuracy is XGBoost.

**Evaluation and Selection of Models (C)**

The following performance indicators were used to assess and compare the models' effectiveness:

- **Accuracy:** This metric determines what proportion of predictions were correct out of all the predictions. The model's accuracy in reducing false positives and in detecting actual positives are shown by the metrics Precision and Recall, respectively.
- **The F1-Score**, which takes accuracy and recall into account, is a comprehensive metric for evaluating model performance.

**The Ensemble Approach**

The following procedures are put into place to improve the final illness prediction's accuracy and reliability: Once the models are loaded, predictions are made using a specific symptom input using the soft voting mechanism. The 42 diseases probability distributions are produced by each model.

- **Averaging Probabilities:** After calculating the probabilities from each model, the final prediction is based on the illness with the greatest overall probability. This ensemble technique enhances accuracy and generates a more reliable final prediction by using the unique characteristics of each model.
- Section E: Working with Online Applications and Users Through the app.py script's Flask-based interface, users may choose up to three symptoms from drop-down menus and input demographic data (gender, age).

**Prediction Process in the Backend**

- A binary vector is generated from the input symptoms to indicate which ones are present.
- Every loaded model uses this symptom vector to estimate the likelihood of sickness; finally, the soft voting ensemble takes all of the results and makes a final prediction.
- The software displays forecasts from all of the models, not just the ensemble prediction, to make sure that everyone can see how decisions are made.

**AI-Generative Patient Health Records**

The system incorporates a generative AI API to improve the user experience and provide deep insights. After predicting an illness, a request is sent to the API using the POST method, followed by the retrieval of additional diagnostic data via the GET method. The purpose of the prompt is to elicit detailed information on the anticipated illness, such as its nature, potential symptoms, causes, therapies, and drugs that may be recommended. For better readability, follow the given formatting guidelines to make sure your headers are legible and your text is organised. A comprehensive summary of the user's health state is provided by the AI-generated data, which is subsequently presented on the homepage in an organised HTML manner.

**Sample Prompt**

Give a comprehensive overview of [Predicted Disease] for a [gender]-year-old [age]. Give details about: Disease overview provided in the description.

- **Factors Contributing:** Typical contributors and potential danger lights. Key symptoms that patients may suffer are included in the symptoms section.
- **Treatments:** typical methods for addressing the problem.
- **Medications:** Medications that are often administered. Guide to Formatting: Sections should be clearly labelled with headings.

Fill out each section thoroughly

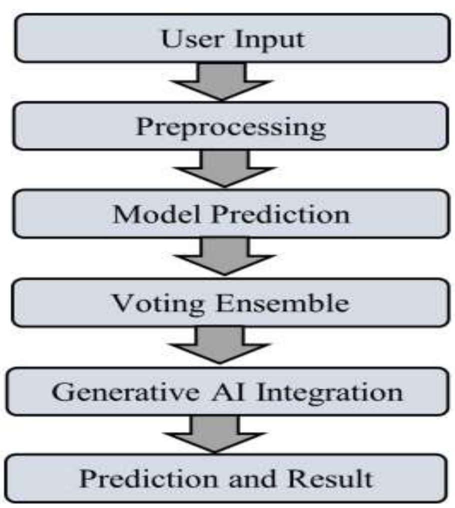


Fig. 2. Block Diagram of Proposed Model

## IV. RESULTS AND DISCUSSION

This section provides a synopsis of the results and analysis from the machine learning models that were used to predict diseases based on symptoms. The models' overall effectiveness is evaluated using a number of performance measures, including accuracy, precision, recall, and F1-score.

Several models were tested, including Decision Tree, Random Forest, Naive Bayes, Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbours (KNN), and XGBoost.

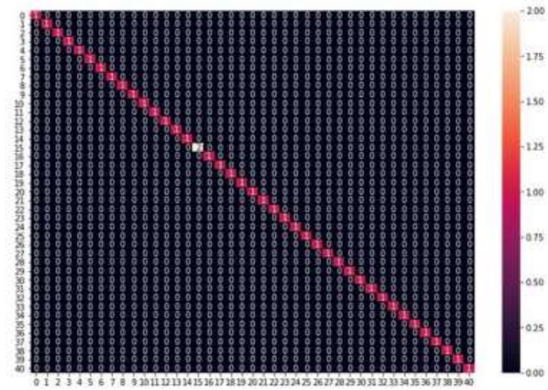

Fig. 3. Confusion Matrix

You can see the results of the classifications in figure 3, which displays the confusion matrices for all the models:

- Diseases that were correctly identified: True Positives (TP).
- Accurately recognised instances without illness are known as True Negatives (TN).
- Misclassified healthy cases as diseased; this is known as a false positive (FP). Misdiagnosed sick patients as healthy ones; this is known as a false negative (FN). The matrices indicate:
- Logistic Regression, KNN, SVM, and Naive Bayes all showed consistently high TP and TN values.
- A larger number of false positives and false negatives caused the Decision Tree model to show reduced accuracy.

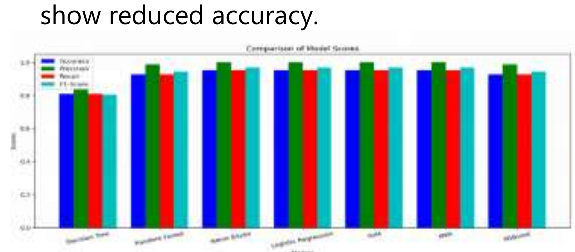

Fig. 4. Bar Chart

Figure 4 is a bar chart comparison that shows how all the models did on each of the four criteria:

- The top-performing methods were Logistic Regression, KNN, SVM, and Naive Bayes, all of which achieved metrics near to or equal to 1.0.
- Random Forest was not far behind the top models, but it still did well.
- The Decision Tree's lower performance was caused by its simpler structure and increased sensitivity to overfitting, even though it achieved respectable results.

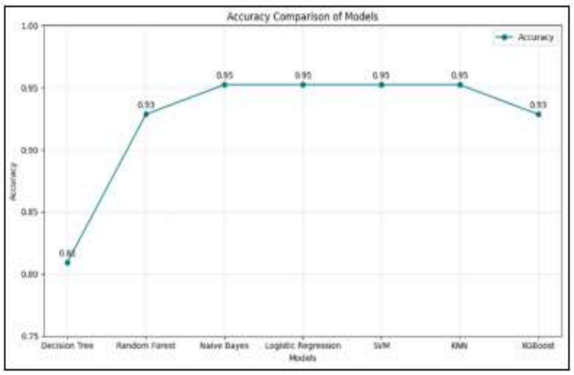


Fig. 5. Accuracy Chart

Figure 5 shows the results of comparing the accuracy of several machine learning models that were used in the suggested system. A few examples of the models include XGBoost, Decision Tree, Random Forest, Naive Bayes, Logistic Regression, and Support Vector Machine (SVM).

- With an accuracy of 81%, the Decision Tree model was the least accurate, suggesting that it used a simpler strategy to handle the information.
- By capitalising on its ensemble character, Random Forest increased accuracy to 93%.

A peak accuracy of 95% was attained by Naive Bayes, Logistic Regression, SVM, and KNN, indicating their exceptional performance in illness prediction tasks.

With an accuracy of 93%, XGBoost was competitive, however it was lower than other models. Optimal performance in this research was achieved by Naive Bayes, Logistic Regression, SVM, and KNN, as seen in this chart, which also shows the efficacy of the various models.

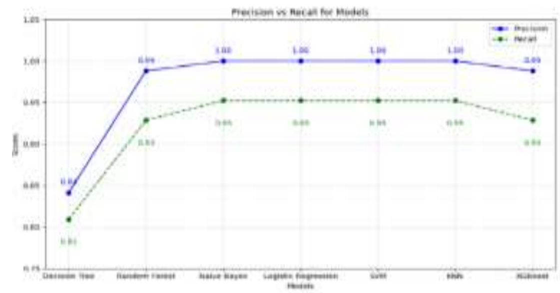


Fig. 6. Precision vs Recall Curve

Figure 6 shows the accuracy-recall curve, which shows how all models compromise on either precision or recall:

- AUC (Area Under Curve) values were high because models like Naive Bayes, Logistic Regression, SVM, and KNN achieved near-perfect recall and accuracy.
- While other models demonstrated more optimal performance, Decision Tree and Random Forest models demonstrated just adequate recall and accuracy.

Overall, the ensemble model performed better than the individual models. This model integrates all the predictions from the separate models using a majority voting technique. The overall accuracy, precision, recall score, and F1-Score were all 96%, 95%, and 96%, respectively, thanks to the ensemble's use of each model's capabilities.

TABLE I
PERFORMANCE METRICS OF MACHINE LEARNING MODELS

| Model | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| Decision Tree | 80.95 | 84.13 | 80.95 | 80.56 |
| Random Forest | 92.86 | 98.81 | 92.86 | 94.44 |
| Naive Bayes | 95.24 | 100.00 | 95.24 | 96.83 |
| Logistic Regression | 95.24 | 100.00 | 95.24 | 96.83 |
| SVM | 95.24 | 100.00 | 95.24 | 96.83 |
| KNN | 95.24 | 100.00 | 95.24 | 96.83 |
| XGBoost | 92.86 | 98.81 | 92.86 | 94.44 |

Based on the results, Naive Bayes, Logistic Regression, Support Vector Machines, and K-Nearest Neighbours are the best models for this dataset. Their capacity to find a happy medium between accuracy and memory makes them ideal for usage in the medical field.

Although interpretable, the Decision Tree needs adjustment to compete with more complex models such as Random Forest and XGBoost. The ensemble model is the most dependable method because it takes the best features of many models and uses them together to provide consistently high accuracy and dependability. This highlights the importance of group techniques while making smart decisions.

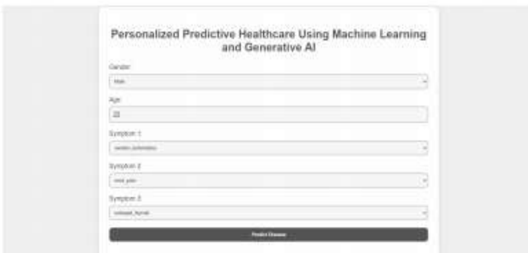


Fig. 7. Input Page

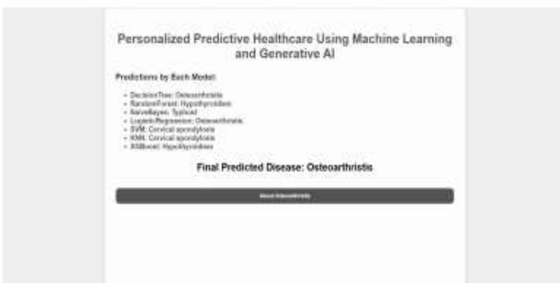


Fig. 8. Prediction Page

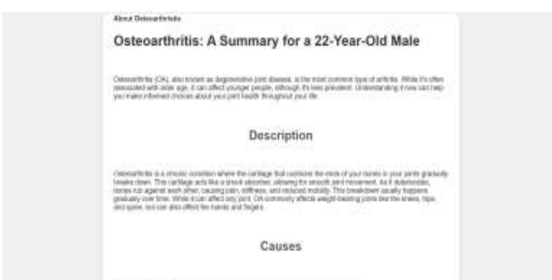


Fig. 9. Detailed Description

Beginning with the input page interface (Figure 7), Figures 7, 8, and 9 show the user journey inside the healthcare system. To make sure everything goes well and is easy to use, this page is the first point of interaction; it asks for basic demographic information like age and gender and lets users choose up to three symptoms. Following this, users will be able to see illness predictions made using their input on the prediction site (Figure 8). This page displays the final consensus result obtained using the voting en semble approach as well as the individual predictions from seven machine learning models.

Finally, a generative AI module creates a descriptive description page (Figure 9) that provides an in-depth account of the projected ailment. By taking the user's age and gender into account, this summary encom passes on information on the illness, including its description, causes, symptoms, cures, and treatment choices. This personalised healthcare knowledge is very helpful.

It is instructive to compare the suggested strategy to well-known machine learning algorithms, as no prior study has integrated generative AI with a voting ensemble model for illness prediction. Support Vector Machines (SVMs), Decision Trees, Naive Bayes, and Random Forest are some of the more traditional models used for healthcare prediction.

Performance and reliability are both enhanced in the suggested system because to the ensemble technique, which integrates many models. On top of that, unlike conventional methods, it provides detailed and personalised insights into sickness. The addition of a generative AI module improves the system's usability, making it easier to use and more accessible. Together, these features make the suggested method stand out as an approach to early illness prediction that is both thorough and user-centric.

## V. CONCLUSION AND FUTURE WORK

**Conclusion**

This work created and assessed an ensemble method to illness prediction in addition to seven separate machine learning models: Decision Tree, Random Forest, Naive Bayes, Logistic Regression, SVM, KNN, and XGBoost.

Among the methods tested, Naive Bayes, Logistic Regression, Support Vector Machines (SVM), and K-Nearest Neighbours (KNN) achieved the highest levels of accuracy, recall, and F1-scores. The most resilient model was the ensemble model, which integrated predictions from all the other models. It had an accuracy of about 96% and strong dependability across all assessment criteria. These results demonstrate the promise of machine learning, and more specifically ensemble approaches, for improving medical diagnosis accuracy.

**Future Work**

To improve the dataset's generalizability and resilience across populations, future work will include adding additional illnesses, demographic variety, and real-world variations. One of the main goals is to make the voting ensemble model as efficient as possible to use in low-resource settings, such mobile and edge devices. To further improve accuracy and efficiency, it is worth looking at deep learning models and finding ways to enhance feature selection.

Creating explainable AI (XAI) tools like SHAP or LIME may significantly increase usability by making predictions more interpretable and building trust with patients. Integrating the technology with real-world medical workflows, including EHR systems, would allow for smooth incorporation with current healthcare procedures. More specifically, considering complex ensemble methods like stacking or weighted voting and tackling the difficulties of identifying uncommon illnesses can enhance prediction performance and, in the long run, diagnostic accuracy.

## REFERENCES

1. S. Dongre, R. Chandra, and S. Agarwal, "MLtoGAI: Semantic Webbased with Machine Learning for Enhanced Disease Prediction and Personalized Recommendations using Generative AI," Indian Institute of Information Technology, Allahabad, India, pp. 1-15, 2024.
2. A. Zolhavarieh, "Tuberculosis classification using SWRL rulebased system," Journal of Health Informatics, vol. 10, no. 3, pp. 102–109,2019.
3. R. Chandra, "Vector-borne Disease (VBD) ontology for vectorborne diseases classification using RDF and SWRL," International Journal of Biomedical Ontology, vol. 6, no. 4, pp. 51–62, 2021.
4. I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative Adversarial Networks," Advances in Neural Information Processing Systems, vol. 27, pp. 1–9, 2014.
5. K. Manubolu and R. Peeriga, "Generative AI in Disease Diagnosis," *Preprints*, vol. 8, no. 2, pp. 55–73, Oct. 2024, doi: 10.20944/preprints202410.1555.v1.
6. S. El Massari, M. Salima, and T. Mebarki, "SWRL and ontology-based inference in oncology," *International Journal of Cancer Informatics*,vol. 14, no. 1, pp. 24–32, 2020.
7. R. Sandh, "Case-based reasoning for dengue symptom classification,"*J. Med. Case Stud.*, vol. 17, no. 2, pp. 87–93, 2022.
8. H. Bensalah, "Bone tumor diagnosis using SWRL-based ontologymodel," Journal of Computational Oncology, vol. 3, no. 5, pp. 45–52,2019.
9. C. Devi, and P. Narayan, "Semantic rule-based dengue classificationusing SWRL," Health Information Science and Systems, vol. 15, no. 2,pp. 112–120, 2023.
10. M. T. Ribeiro, S. Singh, and C. Guestrin, "LIME: Local Interpretable Model-Agnostic Explanations," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and DataMining*, vol. 22, pp. 1135–1144, 2016.
11. S. Dongre, R. Chandra, and S. Agarwal, "ChatGPT in Explainable AI for healthcare," AI in Medicine Journal, vol. 9, no. 3, pp. 45–57, 2024.
12. A. Devi, and K. P. Kumar, "Semantic rule-based dengue classification using SWRL," Journal of Medical Informatics, vol. 10, no. 6, pp. 134–142, 2022.
13. R. Sandh, and T. J. Walsh, "Ontology-driven decision-making for dengue classification," IEEE Transactions on Health Informatics, vol. 18, no. 3, pp. 54–63, 2021.
14. K. Rizwan, et al., "Advanced medical imaging techniques using GANs," International Journal of Imaging Systems and Technology, vol. 20, no. 4, pp. 350–362, 2022.

15. Y. Gong, et al., "Medical imaging quality improvements using GANs," IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 2, pp.67–78, 2021.
16. J. Marie, and C. B. Lee, "Segmentation algorithms in MRI enhancement," Journal of Magnetic Resonance Imaging, vol. 14, no. 5, pp. 67–74, 2018.
17. V. Zlokolica, and P. Vesna, "Image segmentation in cardiovascular diagnostics," Cardiovascular Imaging Journal, vol. 11, no. 1, pp. 25–31, 2019.
18. L. Haubold, S. Gong, and A. Levine, "Deep learning-based image reconstruction in CT," IEEE Transactions on Medical Imaging, vol. 39, no. 4, pp. 750–760, 2020.
19. Glass AI Team, "LLM for clinical decision support," Journal of Artificial Intelligence in Medicine, vol. 12, no. 4, pp. 180–189, 2024.
20. DiseaseDiagNet Research Group, "Diagnostic table generation for patient assessment," Journal of Clinical Decision Support, vol. 9, no. 3, pp. 132–145, 2023