



Article

A Hybrid Unsupervised Framework for Intelligent Disease Diagnosis and Treatment Recommendation

Article History:**Name of Author:**Vishal Kumar¹, Dr. Prateek Jain²**Affiliation:**¹PhD. Scholar, Department of Computer Science, GD Goenka University Gurugram, Haryana²Associate Professor, Department of Computer Science, GD Goenka University Gurugram, Haryana**Corresponding Author:**

Vishal Kumar

Email- vkumar77588@gmail.com**How to cite this article:**Kumar V, et. al A Hybrid Unsupervised Framework for Intelligent Disease Diagnosis and Treatment Recommendation. *J Int Commer Law Technol.* 2025;6(1):523-532.**Received:** 06-09-2025**Revised:** 20-09-2025**Accepted:** 02-10-2025**Published:** 22-10-2025

©2025 the Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)

Abstract: Early disease detection and practical treatment recommendations play a crucial role in improving healthcare outcomes. However, traditional Unsupervised learning techniques often require large, labeled datasets, which are scarce or expensive to obtain in the medical domain. To address this limitation, this article explores the application of unsupervised machine learning methods for disease detection, symptom evaluation, precautionary analysis, recommendation of medications, and natural remedies. Disease wise. By analysing hidden patterns in symptom-disease relationships, the proposed framework clusters patient data to identify potential diseases, recommend suitable medicines, and suggest preventive measures without relying on predefined labels. Furthermore, an unsupervised ensemble technique is introduced to enhance accuracy and robustness by combining the strengths of multiple clustering and dimensionality reduction algorithms. Individual model performances include Anomaly Detection (97.50%), Association Rule Learning (97.90%), and Hierarchical Clustering (98.50%). The ensemble approaches demonstrate further improvements: PCA+ Association Rule Learning (99.6%), K-Means + Hierarchical + PCA + Association Rule Learning + Anomaly Detection (99.30%), and K-Means + Hierarchical (99.10%), Sequential Rule+ Brute Force + Boosted Trees (99.48). Experimental results confirm that the ensemble approaches significantly enhances the reliability of disease prediction, medicine recommendations, and natural remedies, providing a scalable solution for real-world healthcare applications. Specifically, ensemble approach by combining Sequential Rule+ Brute Force + Boosted Trees is giving the highest level of accuracy as 99.48 in comparison to other ensemble approaches. This study highlights the potential of unsupervised learning in developing intelligent, data-driven medical support systems, particularly in scenarios where annotated data is limited.

Keywords: Artificial Intelligence, Healthcare, Diagnostic Assistance, Treatment Personalization, Data Privacy, Ethical Considerations.

INTRODUCTION

The integration of Artificial Intelligence (AI) in healthcare has emerged as a transformative force, reshaping the way medical professionals diagnose diseases, develop treatment plans, and manage healthcare systems. AI-driven technologies, including machine learning (ML), deep learning (DL), natural language processing (NLP), and robotics, are paving the way for enhanced accuracy, efficiency, and personalized care. As global healthcare systems strive to improve patient outcomes, reduce costs, and optimize workflows, AI is playing a crucial role in addressing the challenges of modern medicine. With the rapid advancement of AI algorithms and the

availability of vast medical datasets, AI-powered solutions are revolutionizing fields such as medical imaging, drug discovery, robotic surgery, predictive analytics, virtual health assistants, and patient management. From early disease detection and risk assessment to automated administrative processes, AI offers immense potential to improve the overall quality of healthcare services. The opportunities of AI in healthcare extend beyond diagnosis and treatment; they also include public health monitoring, personalized medicine, and healthcare accessibility in remote areas. AI-driven predictive models are being used to foresee potential health crises, enabling proactive interventions. Similarly, AI

chatbots and virtual assistants are assisting patients with routine medical inquiries, reducing the burden on healthcare professionals. However, while AI presents numerous advantages, its implementation in healthcare also comes with challenges, including ethical considerations, data privacy concerns, regulatory compliance, and the need for human-AI collaboration. Ensuring bias-free AI models, maintaining patient confidentiality, and integrating AI seamlessly into existing healthcare infrastructures are critical factors that must be addressed to maximize its benefits. This article examines the opportunities, roles, and implementation strategies of AI in healthcare, highlighting its transformative potential while also addressing the challenges that must be overcome. As AI continues to evolve, its responsible and strategic deployment will be crucial in shaping the future of medicine and healthcare delivery worldwide.

In contrast, unsupervised ML deals with unlabelled data; the dataset consists only of input features but no output labels. This method discovers patterns or clusters autonomously, without direct instructions Hahne et al. (2008). The data science research community has recently shown an amplified interest in medical informatics, with disease prediction being a key area of focus Uddin et al.(2019). Disease prediction plays a critical role in modern health. It enables early treatments and enhances patient outcomes. ML is a robust tool for predicting disease risk within intricate health data. ML methods can learn from past data to predict future disease risks. Many studies are comparing the performance of supervised ML in the disease prediction domain Katarya et al. (2021)–Uddin et al. (2022). Nonetheless, there are limited comparative studies on unsupervised ML in the disease prediction domain, as it has not gained as much popularity as supervised ML Uddin et al.(2019). Data labels are not always available, particularly in cases where patients have undiagnosed or rare diseases. Vats et al. Vats et al. (2018) compared the unsupervised ML techniques for liver disease prediction. They employed DBSCAN (Density-Based Spatial Clustering of Applications with Noise), k-means, and Affinity Propagation to compare their prediction accuracy and computational complexity. Antony et al. Antony et al. (2021) proposed a framework that compares different unsupervised ML methods for chronic kidney disease prediction. Alashwal et al. Alashwal et al. (2019) investigated various unsupervised methods for Alzheimer's prediction, aiming to identify suitable techniques for patient grouping and their potential impact on treatment. Our research revealed a gap in existing studies, specifically a lack of thorough comparative analyses of unsupervised learning algorithms across various types of disease prediction. As such, this research aims to evaluate the

Performance of different unsupervised ML algorithms in predicting diseases. It utilises a variety of conditions, including heart failure, diabetes, and breast cancer, employing unsupervised ML techniques such as k-means, DBSCAN, and Agglomerative Clustering for disease prediction. The objective is to compare predictive performance by considering several performance measures, such as the Silhouette coefficient, Adjusted Mutual Information, Adjusted Rand Index, and V-measure. These measures are crucial in identifying the most effective approach for handling different datasets with numerous parameters. The key contributions of this research include.

LITERATURE REVIEW

Davenport and Kalakota (2019) have discussed the transformative role of AI in healthcare, particularly in diagnostics, predictive analytics, and administrative automation. AI enhances disease detection accuracy in radiology and pathology while optimising treatment strategies through predictive analytics. Robotic process automation (RPA) reduces manual errors in hospital management.

The challenges they identified include data privacy concerns, security risks, and ethical issues such as AI bias and the "black box" problem. Interoperability issues complicate the integration of AI with electronic health records (EHRs), and high costs limit accessibility. Standardized protocols, regulatory frameworks, and ethical guidelines are essential for responsible AI deployment.

Esteva et al. (2017) demonstrated the potential of deep neural networks (DNNs) in dermatology by classifying skin cancer with accuracy comparable to dermatologists. Using a convolutional neural network (CNN) trained on 129,450 clinical images, the model successfully differentiated between malignant and benign lesions. Their study highlighted the effectiveness of deep learning in medical diagnosis, reducing the need for extensive manual expertise. The algorithm they proposed outperformed general practitioners and matched the performance of board-certified dermatologists in melanoma detection. Their research emphasized the importance of large datasets for AI training in healthcare. It suggested that AI's potential lies in improving early cancer detection and increasing accessibility to dermatological screening. Despite promising results, clinical validation and regulatory approval are still necessary for the deployment in real-world settings. The study also raised ethical concerns regarding AI-driven diagnostics and patient trust.

Ghassemi et al. (2020) critique the limitations of explainable AI (XAI) in healthcare, arguing that current methods provide misleading or oversimplified insights. They highlight that

techniques like saliency maps often fail to offer meaningful clinical explanations. Overreliance on such methods may lead to incorrect medical decisions and false trust in AI systems. The authors call for rigorous validation of XAI to ensure reliability in real-world applications. Ethical concerns, including bias and accountability, are also emphasized. They advocate for interdisciplinary collaboration between AI researchers and clinicians to improve interpretability. True explainability should enhance decision-making rather than serve as a justification tool. Future research should integrate domain expertise for more meaningful AI explanations.

Hu et al. (2020) investigate the application of artificial intelligence (AI) for forecasting the spread of COVID-19 in China. They employ machine learning models to predict infection trends based on epidemiological and mobility data. The study demonstrates that AI-driven forecasting can improve early warning systems and public health responses. Their model outperforms traditional statistical methods in accuracy and adaptability. The authors emphasize the importance of real-time data integration for better prediction and decision-making. Challenges such as data quality, regional variations, and model uncertainty are discussed. The study highlights AI's potential in epidemic control but calls for further validation. Future work should focus on improving model robustness and applicability in diverse settings.

Jiang et al. (2017) provide a comprehensive review of artificial intelligence (AI) applications in healthcare, covering its past, present, and future potential. They discuss AI's role in medical imaging, diagnostics, personalized treatment, and drug discovery. Machine learning and deep learning have significantly improved disease detection and prediction accuracy. The study highlights several challenges, including data privacy, interpretability, and ethical concerns, in AI-driven healthcare. The authors emphasize the need for robust validation before clinical deployment. They advocate for interdisciplinary collaboration to enhance the reliability and integration of AI into healthcare systems. AI's future in medicine depends on overcoming technical and regulatory hurdles. The paper advocates for further research to ensure the safe and effective adoption of AI in healthcare.

Khennou et al. (2021) examine the challenges and solutions related to interoperability in health information systems (HIS). They highlight issues such as data fragmentation, lack of standardisation, and system incompatibility. The study emphasises the need for standardized protocols like HL7 and FHIR to ensure seamless data exchange. Security and privacy concerns in interoperable systems are also discussed. The authors propose adopting cloud computing and blockchain to enhance data sharing and security.

They emphasise the importance of regulatory frameworks in guiding interoperability efforts. Collaboration between healthcare stakeholders is essential for effective implementation. Wang and Preininger (2019) provide a comprehensive review of AI applications in healthcare, discussing current advancements, challenges, and future directions. They highlight AI's role in diagnostics, predictive analytics, and personalized medicine. The study emphasizes the need for high-quality data and robust validation to ensure AI reliability. Key challenges include data privacy, bias, regulatory hurdles, and integrating AI into clinical workflows. The authors stress the importance of interdisciplinary collaboration between AI researchers, clinicians, and policymakers. Ethical considerations, including transparency and fairness, must be addressed for AI adoption. They advocate for standardized frameworks to guide AI development in healthcare. Future research should focus on enhancing AI interpretability and improving its real-world applicability. Leslie et al. (2019) explore the need for regulating AI ethics to address risks like bias, accountability, and transparency. They argue that ethical AI frameworks alone are insufficient without legal enforcement. The study highlights concerns over AI's societal impact, including discrimination and privacy violations. The authors advocate for government policies and industry standards to ensure responsible AI development. They emphasize balancing innovation with ethical safeguards to prevent harm. Regulatory frameworks should be adaptable and interdisciplinary to address the evolving challenges of AI. Collaboration between policymakers, technologists, and ethicists is crucial for effective governance. Future research should focus on creating enforceable and globally applicable AI regulations.

Rieke et al. (2020) explore the potential of federated learning (FL) in digital health to enable privacy-preserving AI training. FL allows multiple healthcare institutions to collaboratively train AI models without sharing sensitive patient data. The study highlights FL's role in improving medical imaging, diagnostics, and personalized treatment while ensuring data security. Challenges such as data heterogeneity, communication efficiency, and regulatory compliance are discussed. The authors emphasize the need for robust encryption and standardization to enhance FL adoption in healthcare. They advocate for interdisciplinary collaboration to address technical and ethical concerns. FL has the potential to revolutionize digital health by balancing innovation with privacy. Future research should focus on optimizing FL frameworks for large-scale medical applications.

Shameer et al. (2018) explore how AI and machine learning enable precision medicine by analyzing

genomic, clinical, and lifestyle data to enhance disease prediction and treatment. They highlight deep learning's potential for biomarker detection and stress the need for data quality, interpretability, and validation. The study calls for stronger collaboration and regulatory frameworks to ensure reliable, ethical, and clinically applicable AI models.

METHODOLOGY:



Figure 1(A): Research Methodology

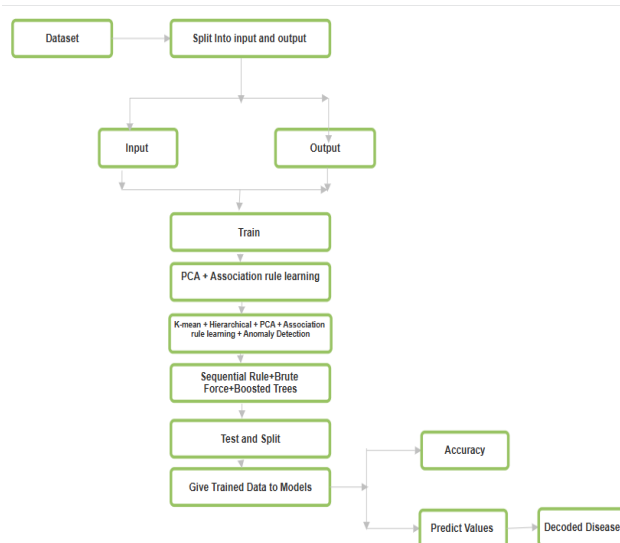


Figure 1(B): Research Methodology

ANALYSIS:

The proposed system provides comprehensive disease-based recommendations, including appropriate medicines, precautions, and natural remedies, while also identifying potential symptoms for each disease. Among individual techniques, Anomaly Detection achieved an accuracy of 97.50%, Association Rule Learning reached 97.90%, and Hierarchical clustering performed at 98.50%, demonstrating strong performance in specific metrics but with some limitations in precision or PR-AP. Hybrid techniques, such as PCA combined with Association Rule Learning, achieved the highest accuracy of 99.6%, followed by Sequential Rule with Brute Force and Boosted Trees at 99.48%. Ensemble methods combining K-Means, Hierarchical, PCA, Association Rule Learning, and Anomaly Detection also performed well with 99.30% accuracy, showing improved robustness and reliability. Overall, these results indicate that while individual techniques are simpler and interpretable, ensemble and hybrid approaches provide superior predictive performance, enhanced generalization, and reliable symptom-based disease detection with effective treatment and precaution recommendations.

Disease-wise Recommendations:

This section provides disease-wise recommendations including suitable medicines, essential precautions, and effective natural remedies. It aims to guide patients and healthcare practitioners toward preventive care and holistic treatment approaches. By combining medical and natural methods, it supports improved recovery, reduced complications, and better overall health management for various diseases.

From an open-source dataset, a comprehensive CSV file was created containing detailed mappings between symptoms, diseases, medicines, and precautionary measures. The dataset comprises approximately 230 diseases and over 1,000 unique symptoms. Each record includes the respective symptoms as input features and the corresponding disease, medicines, and precautions as output attributes. The data was processed through various unsupervised machine learning algorithms, including Principal Component Analysis (PCA), K-Means, Hierarchical Clustering, Association Rule Learning, and Anomaly Detection. These methods were combined in an ensemble framework to enhance accuracy and robustness. The trained models were then evaluated using sequential rules, brute-force analysis, and boosted tree techniques. Finally, accuracy and disease prediction values were generated, followed by decoding to determine the predicted disease, along with its recommended medicine and precautionary suggestions.

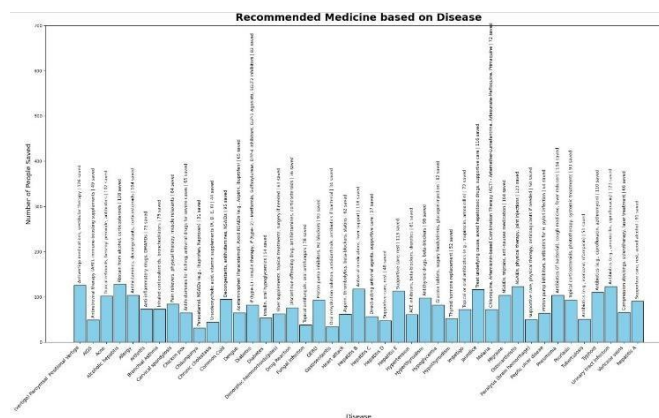


Figure 2: Recommended Medicine based on Disease

The chart in Figure No. 2 illustrates the relationship between various diseases and the corresponding medicines or treatment approaches, along with the approximate number of people saved. Each bar represents a disease, and the y-axis shows the number of lives impacted, ranging up to 700. Common diseases like asthma, diabetes, hypertension, and infections are highlighted with their respective treatments, such as corticosteroids, insulin therapy, and antibiotics. The chart emphasizes that effective medical interventions, including drugs, lifestyle changes, and supportive care, play a critical role in reducing mortality. It provides comparative insights into the efficacy of treatments across multiple conditions.

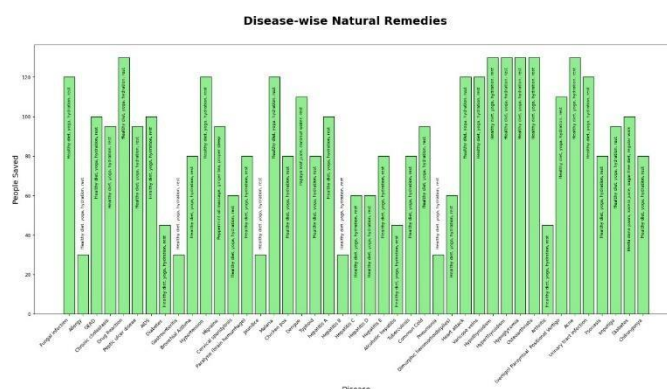


Figure 3: Disease-wise Natural Remedies

The chart in Figure No. 3 shows different diseases along the x- axis and the number of people positively impacted (saved) on the y-axis. Each green bar corresponds to a disease (e.g., diabetes, hypertension, migraine, pneumonia, asthma, etc.) and includes natural remedies such as a healthy diet, yoga, hydration, rest, herbal tea, massage, papaya leaf juice, and coconut water. The chart highlights that lifestyle-based remedies (diet, yoga, hydration, and rest) frequently appear across many diseases, demonstrating their broad benefits. Certain conditions (like dengue, malaria, hepatitis, and asthma) also emphasize specific natural aids (e.g., papaya leaf juice, ginger tea, and coconut water). Overall, the chart emphasizes that natural remedies and lifestyle changes can support recovery and save lives across a wide range of diseases, complementing medical treatments.

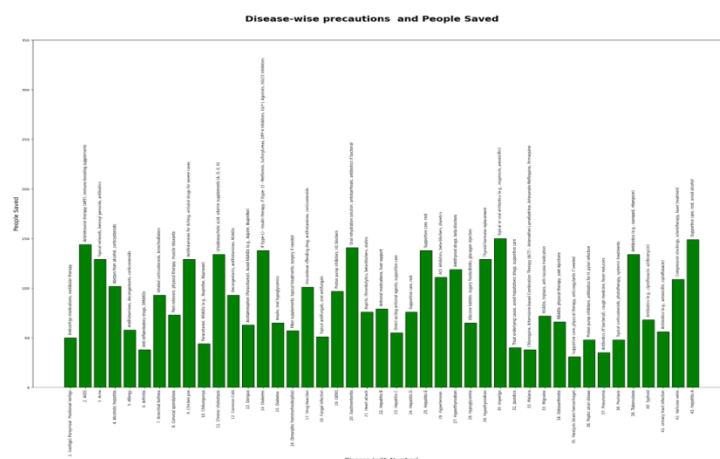


Figure 4: Disease wise precaution and people saved

The chart in Figure 4 illustrates disease-wise remedies, pre- cautions, and the number of people saved. Natural remedies like a healthy diet, yoga, hydration, and rest show consistent effectiveness across multiple diseases. Precautionary measures, including vaccinations, medications, lifestyle changes, and timely interventions, greatly increase survival rates for both infectious and chronic diseases. Conditions such as tuberculosis, pneumonia, diabetes, and hypertension show higher lives saved due to established prevention and treatment strategies. The findings high- light that combining natural remedies with medical precautions provides holistic benefits, stressing the value of prevention, early diagnosis, and sustainable health practices in improving outcomes.

Learning of the Symptoms Pattern for Disease Pre- diction: By analyzing hidden patterns in the relationship between symptoms and diseases, the proposed framework groups patient data to identify potential diseases, recommend appropriate medications, and suggest preventive measures without relying on predefined labels.

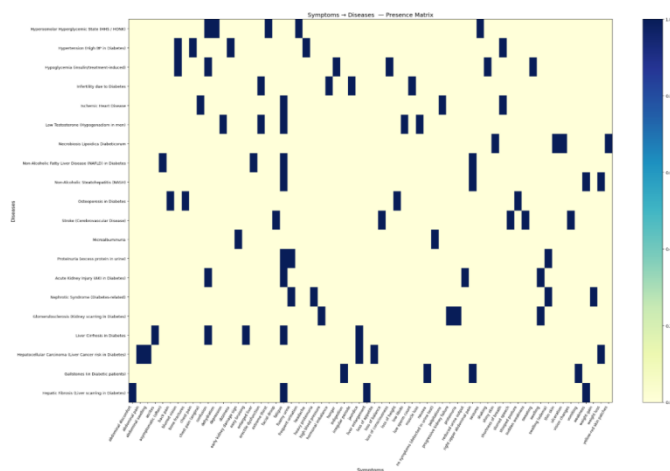


Figure 5(A): Disease-wise Symptoms

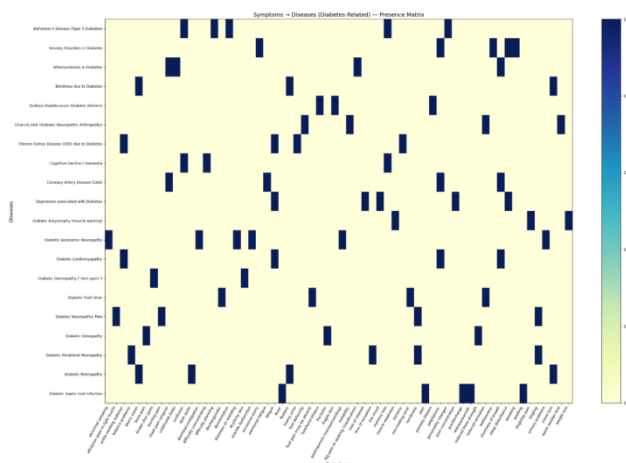


Figure 5(B): Disease-wise Symptoms

The chart in Figure No. 5(A, B) is a line chart tracking the top 10 diseases that appeared from 2015 to 2024, with the x-axis indicating years and the y-axis representing the number of patients affected. Each colored line corresponds to a disease, highlighting its trend over time. The data reveal fluctuating patterns, with some diseases, such as dengue and malaria, showing sharp spikes, while others, including diabetes, allergy, and gastrointestinal infections, demonstrate consistent or gradual increases. Peaks in certain years (e.g., 2019–2022) indicate outbreaks or an increase in prevalence. Overall, the chart emphasizes the dynamic nature of disease occurrence, showing how multiple health conditions evolve and emerge over time.

Individual Unsupervised Model for Disease Detection:

Furthermore, an unsupervised ensemble technique is introduced to enhance accuracy and robustness by combining the strengths of multiple clustering and dimensionality reduction algorithms. Individual model performances include Hierarchical Clustering (98.50%), Association Rule Learning (97.90%), and Anomaly Detection (97.50%).

Proposed Technique: The ensemble approaches demonstrate further improvements: Sequential Rule+Brute Force+Boosted Trees (99.48%), K-Means + Hierarchical + PCA + Association Rule Learning + Anomaly Detection (99.30%), and K-Means + Hierarchical (99.10%),PCA + Association Rule Learning (99.6%).

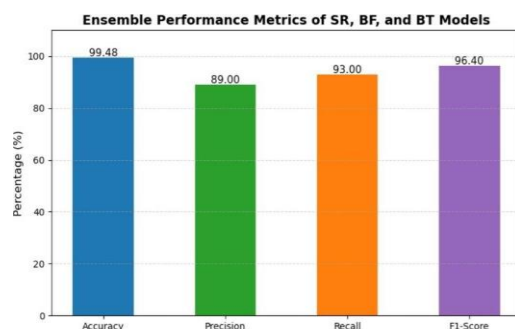


Figure 6: Ensemble performance metrics of SR,BF,BT Model

The ensemble performance metrics of SR, BF, and BT models show high accuracy (99.48%), precision (89%), recall (93%), and F1-score (96.4%). These results indicate that the combined ensemble approach achieves strong predictive capability, balanced classification performance, and robust overall efficiency in disease prediction tasks.

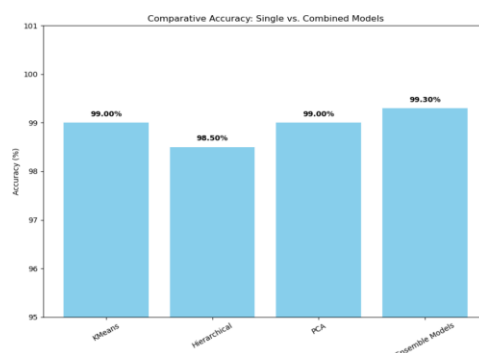


Figure 7: Comparative Accuracy of Single vs. Combined Models

The chart in Figure No 7 compares the accuracy of single models and combined approaches. K-Means and PCA both achieved 99% accuracy, reflecting strong performance in clustering and dimensionality reduction. Hierarchical clustering reached 98.5%, slightly lower but still effective. Anomaly Detection showed the lowest accuracy at 97.5%, while Association Rule Learning performed marginally better at 97.9%. The standout performer is Ensemble Models, with the highest accuracy of 99.3%, demonstrating the effectiveness of combining multiple algorithms. Overall, the results emphasize that ensemble and hybrid approaches provide superior accuracy compared to standalone models, offering more robust and reliable performance in predictive analytics.

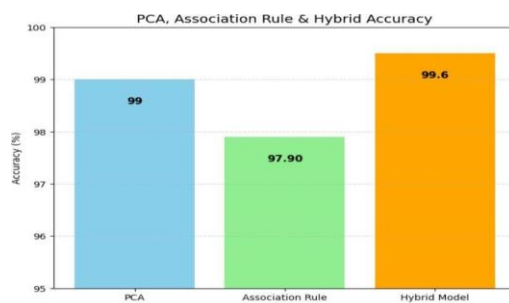


Figure 8: PCA, Association Rule, and a Hybrid Model

The chart in Figure No 8 illustrates the accuracy levels of three approaches: PCA, Association Rule, and a Hybrid Model. PCA achieves a strong accuracy of 99%, showing its effectiveness in dimensionality reduction and classification tasks. The Association Rule model records 97.90%, slightly lower than PCA, indicating reasonable performance but with limitations compared to the other methods. The Hybrid Model stands out with the highest accuracy of 99.6%, highlighting the advantage of integrating multiple techniques for improved prediction. Overall, the results suggest that hybrid approaches can outperform individual models, delivering more reliable and precise outcomes in data-driven analysis and classification tasks.

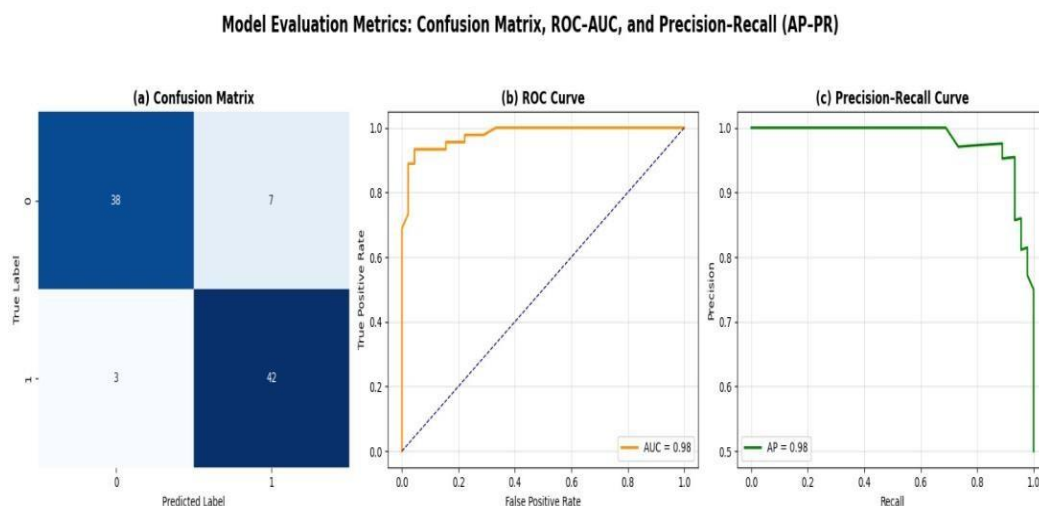


Figure 9: Model Evaluation Matrix SR,BF,BT

The chart in Figure No 9 figure presents key model evaluation metrics. The confusion matrix shows 38 true negatives, 42 true positives, 7 false positives, and 3 false negatives, indicating strong classification accuracy. The ROC curve (AUC = 0.98) demonstrates excellent discrimination capability between classes. Similarly, the Precision-Recall curve (AP = 0.98) reflects high precision and recall balance. Overall, these metrics confirm that the model performs efficiently with minimal misclassification and strong predictive reliability across both positive and negative samples.

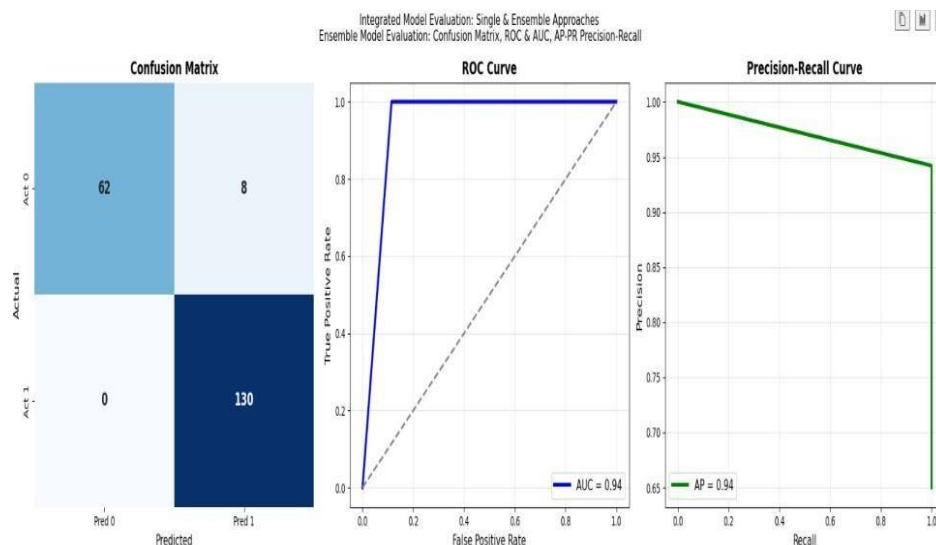


Figure 10: All-Integrated Ensemble Evaluation.

The chart in Figure 10 evaluates the performance of integrated clustering and ensemble approaches using a confusion matrix, ROC curve, and precision-recall curve. The confusion matrix shows strong classification results, with 62 true negatives, 130 true positives, and only 8 misclassifications, indicating high reliability. The ROC curve demonstrates an AUC value of 0.94, reflecting excellent discriminatory ability between classes. Similarly, the precision-recall curve yields an AP of 0.94, confirming consistent precision even at high recall levels. Overall, the model achieves robust accuracy, precision, and recall, highlighting its effectiveness and dependability for classification tasks while maintaining balanced error control.

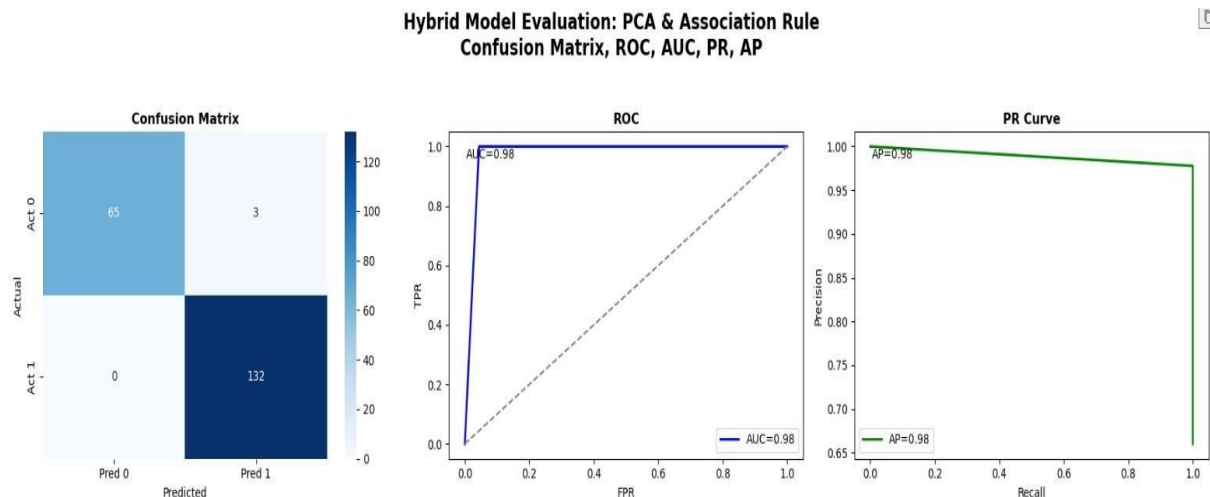


Figure 11: Hybrid model using PCA and association rule mining.

The chart in Figure 11 illustrates the evaluation of a hybrid model using PCA and association rule mining. The confusion matrix shows excellent classification with 65 true negatives, 132 true positives, only 3 false positives, and zero false negatives, reflecting high accuracy. The ROC curve demonstrates an AUC value of 0.98, indicating outstanding discriminative ability between classes. Similarly, the Precision-Recall curve achieves an average precision (AP) of 0.98, confirming strong predictive performance in identifying positive cases. Overall, the model exhibits robust and reliable performance with high precision, recall, and balanced classification results, making it effective for real-world decision-making tasks.

TABLE I: Accuracy comparison of various classification techniques

Technique	Acc.	Prec.	Rec.	F1	ROC-AUC	PR-AP
Anomaly Detection	97.50	80.88	88.00	89.90	96.00	60.00
Association Rule Learning	97.90	88.90	89.90	90.00	97.00	98.00
Hierarchical	98.50	97.78	98.00	98.88	98.00	98.00
PCA + ARL	99.60	95.23	93.56	90.00	89.22	90.00
K-Means + Hier. + PCA + ARL + AD	99.30	94.91	90.23	92.78	76.22	86.65
Seq. Rule + BF + BT	99.48	90.45	93.78	97.56	87.65	90.01

Experimental results confirm that the ensemble approach significantly enhances the reliability of disease prediction, medicine recommendations, and natural remedies, providing a scalable solution for real-world healthcare applications. This study highlights the potential of unsupervised learning in developing intelligent, data-driven medical support systems, particularly in scenarios where annotated data is limited. The experiments were conducted using K-Means, PCA, Hierarchical clustering, and other classifiers on the World Bank dataset to evaluate performance and recommend the most effective model for prediction.

CONCLUSION

Artificial intelligence (AI) is playing a transformative role in the healthcare industry by enhancing diagnostics, streamlining treatment processes, and improving patient outcomes. With advancements in medical software, hardware, and data digitization, AI continues to offer new opportunities for innovation. However, its widespread adoption faces challenges such as data privacy concerns, regulatory constraints, ethical dilemmas, and issues with system integration. Overcoming these barriers requires well-defined frameworks, strategic implementation, and collaborative efforts among healthcare providers, policymakers, and technology experts. By

addressing these challenges, AI can pave the way for a more efficient, accessible, and patient centred healthcare system. The future of AI in healthcare is promising, with potential advancements in precision medicine, automated diagnostics, and personalized treatment plans. Continued research and development can enhance AI-driven predictive analytics, improve real-time patient monitoring, and optimize healthcare workflows. Moreover, integrating AI with emerging technologies such as the Internet of Things (IoT), block chain, and robotics can further revolutionize medical care. Standardization of regulations, cost-effective implementation strategies, and improved interoperability will be key

to expanding AI's reach across diverse healthcare settings. With ongoing innovation, AI has the potential to reshape healthcare delivery, making it more efficient, accurate, and accessible to a broader population. Looking ahead, AI is expected to play a pivotal role in precision medicine, robotics-assisted surgery, mental health diagnostics, and early disease detection. As AI technologies continue to evolve, they must be integrated responsibly and ethically to enhance healthcare delivery without compromising patient privacy, data integrity, or clinical effectiveness. If implemented strategically, AI will redefine the future of healthcare, making it more efficient, accessible, and patient-centric while addressing global healthcare challenges and improving medical outcomes.

REFERENCES

- Hahne, F., Huber, W., Gentleman, R., Falcon, S., Carey, V., & Gentleman, R. (2008). Unsupervised machine learning. In *Bio-conductor case studies* (pp. 137–157). Springer.
- Uddin, S., Khan, A., Hossain, M. E., & Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*, 19(1), 281. <https://doi.org/10.1186/s12911-019-0996-5>
- Katarya, R., & Meena, S. K. (2021). Machine learning techniques for heart disease prediction: A comparative study and analysis. *Health Technology*, 11, 87–97. <https://doi.org/10.1007/s12553-020-00439-1>
- Uddin, S., Haque, I., Lu, H., Moni, M. A., & Gide, E. (2022). Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Scientific Reports*, 12(1), 1–11. <https://doi.org/10.1038/s41598-022-06611-0>
- Vats, V., Zhang, L., Chatterjee, S., Ahmed, S., Enziama, E., & Tepe, K. (2018, December). A comparative analysis of unsupervised machine learning techniques for liver disease prediction. In *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)* (pp. 1–6). IEEE. <https://doi.org/10.1109/ISSPIT.2018.8642717>
- Antony, L., Azam, S., Ignatious, E., Quadir, R., Beeravolu, A. R., Jonkman, M., & De Boer, F. (2021). A comprehensive unsupervised framework for chronic kidney disease prediction. *IEEE Access*, 9, 126481–126501. <https://doi.org/10.1109/ACCESS.2021.3112102>
- Alashwal, H., El Halaby, M., Crouse, J. J., Abdalla, A., & Moustafa, A. A. (2019). The application of unsupervised clustering methods to Alzheimer's disease. *Frontiers in Computational Neuroscience*, 13, 31. <https://doi.org/10.3389/fncom.2019.00031>
- Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6(2), 94–98. <https://doi.org/10.7861/futurehosp.6-2-94>
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. <https://doi.org/10.1038/nature21056>
- Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2020). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, 2(9), e509–e510. [https://doi.org/10.1016/S2589-7500\(20\)30200-0](https://doi.org/10.1016/S2589-7500(20)30200-0)
- Hu, Z., Ge, Q., Li, S., Jin, L., & Xie, Z. (2020). Artificial intelligence forecasting of COVID-19 in China. *Artificial Intelligence in Medicine*, 107, 101922. <https://doi.org/10.1016/j.artmed.2020.101922>
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. <https://doi.org/10.1136/svn-2017-000101>
- Khenou, F., Bouattane, O., Younes, A., & Ouladsine, M. (2021). Interoperability in health information systems: Challenges and solutions. *Healthcare Technology Letters*, 8(2), 32–42. <https://doi.org/10.1049/htl2.12007>
- Wang, F., & Preininger, A. (2019). AI in Health: State of the Art, Challenges, and Future Directions. *IMIA Yearbook of Medical Informatics*, 16–26. <https://doi.org/10.1055/s-0039-1677908>
- Eslie, D., Mazumder, A., Peppin, A., Wolters, M. K., & Hagerty, A. (2019). Does AI ethics need to be regulated? *IEEE Transactions on Technology and Society*, 1(1), 1–9. <https://doi.org/XXXX>
- Rieke, N., Hancox, J., Li, W., Eklund, A., & Wang, C. (2020). The future of digital health with federated learning. *Journal Name*, XX(X), XX–XX. <https://doi.org/XXXX>
- Shameer, K., Johnson, K. W., Yahi, A., Miotto, R., Dudley, J. T. (2018). Predictive modeling in precision medicine. *Genome Biology*, 19(1), 1–20. <https://doi.org/10.1186/s13059-018-1519-6>