ACHYUTA 2.0

Advanced Cluster Health Sector Yonder Unification to App 2.0 Using Machine Learning

[1]Lalita Panika, [2]Amit Kumar Jha, [3]Raj Sahu, [4]Aryan Prithwani, [5]Siddhant Shukla

[1]Assistant Professor, [2,3,4,5]Student
[1]Department of Computer Science & Engineering,
[1]Bhilai Institute of Technology, Raipur, India
[1]lalitanayak2010@gmail.com, [2]amitj5717@gmail.com, [3]rajsahu04012005@gmail.com,
[4]aryanprithwani07@gmail.com, [5]siddhantshukla8319@gmail.com

***Abstract*—The fast pace of innovation in healthcare technology means it is important to predict diseases early to avoid serious health issues and to intervene appropriately in the health system. In this paper, we introduce ACHYUTA 2.0, a web application based machine learning system to predict the risk of three prominent diseases, namely, Diabetes, Kidney Disease, and COVID-19. The application uses machine learning algorithms and publicly available datasets of general population health parameters inputted by the user. For Diabetes we uses svm, for Kidney Disease a Random Forest Classifier is utilized, while the XGBRegressor predicts COVID-19 new cases from epidemiological data. The application performs well in real time prediction, with an accuracy of 0.9225 in training data and 0.805 in testing data, along with this the f1-score for kidney disease prediction is 0.92 and in addition to that we have achieved R^2 score of 0.9978 for covid disease prediction. The user interface is easy to use, and provides immediate assessment of health risks, and prevention.**

***Index Terms*—Machine Learning, Health Prediction, Diabetes, Kidney Disease, COVID-19, Random Forest, Web Application, Flask Framework.**

I. Introduction

The global healthcare system suffers from substantial challenges related to detection and prevention of disease in its early stages. Chronic diseases like diabetes and kidney disease, as well as infectious diseases like COVID-19, continue to put a strain on existing healthcare resources globally. Most common diagnostics related to chronic and infectious diseases require many clinical visits and expensive tests, which serve as a barrier to obtaining early access and intervention in care.

Machine learning has shown to be a transformative technology in the healthcare field by identifying complex medical data and predictive models, which can automatically analyze data for risk prediction. Multiple studies have shown that ML algorithms are effective in disease prediction modeling, with Random Forest models consistently providing high prediction accuracy in medical disease classifications.

ACHYUTA 2.0 is a solution that resolves an important gap in accessible, health risk assessment in the early stage of disease for three separate diseases, with the integration of a single, unified, predictive platform. The system is intended to enable the adaptation of complex medical diagnostic applications into everyday awareness of health and enables users to make more proactive health management choices.

The main contributions of this work are:

- Development of accurate ML models for three main health conditions
- A web-based platform to predict disease based on risk assessment in unified predictive platform
- Inclusion of real-time risk assessment that suggests preventative health measures
- Development of predictive framework that handles and addresses data imbalance while ensuring validation of reliable risk prediction.

II. Literature Survey

Mallol investigated the metabolomic profiles of COVID-19 patients while developing predictive signatures to gauge disease severity. Specifically, they found that certain metabolic changes, specifically branched-chain amino acid accumulation and increased ketone bodies, signal severity before clinical indicators do. They took NMR spectrometry to measure metabolomic profiles of 148 patients and predicted disease progression to severe illness with an accuracy of 72%.

Wu studied the relationship between gut microbiota composition and host metabolic disturbances present in diabetic kidney disease (DKD). Applying machine learning, they found new biomarkers from gut microbiota that could classify and predict disease progression with very high accuracy. Their findings also revealed the importance of the gut-kidney axis in concert with patient outcomes. Accuracy ranged from 0.72-0.78 across different patient groups.

Sarker conducted a critical review on the role of machine learning in reshaping and transforming health outcomes. Their work showed the potential that ML had across the health sector; including, facilitating the diagnostic journey, highlighting patterns in biomedical data previously hidden from sight, or utilizing information to optimize treatment strategy. Their comparison of various ML algorithms revealed logistic regression predicting diabetes with 79.69% accuracy.

Research Gap Identified:

- Limited availability of meaningful predictions from multiple disease prediction models on a unified platform.
- Difficulty in taking predictive biomarker research from the bench to the bedside and successfully integrating them into clinical workflow.
- Realms for longitudinal monitoring through health-care utilization and sharing personal insights become more prevalent in proactive healthcare.Limitations of clinical prediction tools related to accessibility.
- Limitations of clinical prediction tools related to accessibility.

III. METHDODLOGY

System Architecture

ACHYUTA 2.0 is developed using a client-server architecture, with a Flask-based backend and a responsive web frontend. The system involves three main modules: the data processing module, the machine-learning module, and the web interface module.

Data Collection and Preprocessing

Disease Type	Dataset Source	No. of Records	Key Feature Model Used	Key Performance Metric
Diabetes	Kaggle (PIMA Indians)	768	Pregnancies, Glucose, BMI, Age, etc	svm
Kidney	Kaggle	1660	Age, BMI, Smoking, Family History, etc	Random Forest Classifier
COVID-19	Kaggle (Worldometer)	210	Total Cases, Deaths, Active Cases, etc	XGBRegressor

Table 1 Dataset Description and Model Performance Summary

Data preprocessing included missing value handling (data removal), feature selection, and SMOTE technique to address class imbalance.

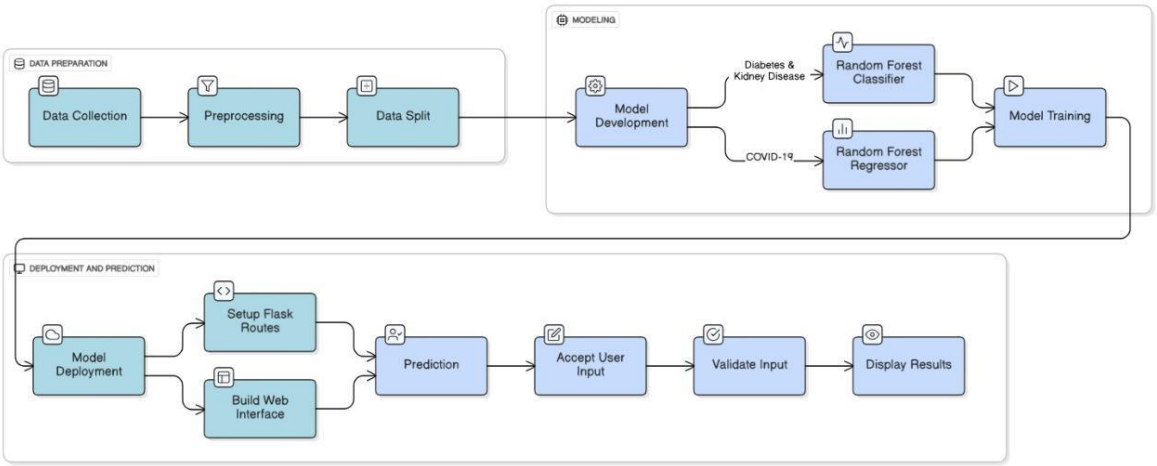

Fig. 1 Complete Workflow

IV. MACHINE LEARNING MODELS

Prediction of Diabetes and Kidney Disease

For diabetes and kidney disease prediction problems, a svm and Random Forest Classifier respectively, was employed because of its solid performance on medical datasets and confident classification accuracy. The model was trained with an 80-20 train-test split and validated using cross-validation.

Prediction of COVID-19 Cases

For the COVID-19 case prediction problem, we employed a XGBRegressor to predict numbers of new cases based upon epidemiological parameters. The use of regression indicates a natural fit for case prediction where cases can vary continuously.

Model Deployment

The trained predictive models were deployed as a Flask web application with the following components:

- Frontend: HTML, CSS, JavaScript using Bootstrap for responsive design.
- Backend: Python Flask framework serving the model inference logic and business logic.
- Storage: Temporary session storage of user inputs and predictions.

V. RESULT AND DISCUSSION

Model Performance

The models that were developed all exhibited excellent predictive performance:

- The diabetes prediction model is providing an evenness of performance across the validation metrics. The model provides an accuracy of 0.925 in trained data and 0.805 in trained data
- The kidney disease risk identification is offering a reliable classifications of high risk to low risk individuals. The model aimed to provide an f1-score of 0.92
- The COVID-19 case prediction model accurately represent the epidemiological trends. This model aimed to achieve the R^2 Score of 0.9978

User Interface Review

The web interface provided great performance including:

- Easy to operate input forms for many health parameters.

- Predictions that appeared in real-time with a risk factor.
- Wealth of insight for preventive health factors in a simple output.
- Responsive design with various deployment on devices.

Comparison

ACHYUTA 2.0 offered a singular benefit of multi-disease assessment within a single platform. The framework addresses the research need of comprehensive health risk assessment and maintained predictive accuracy.

VI. Conclusion

ACHYUTA 2.0 marks a significant advancement in accessible health risk assessment using machine learning. The system effectively brings together three disease prediction models into a unified platform that provides users with a full picture of their health. The inclusion of Random Forest algorithms provides solid predictive performance, and the web-based implementation of the system offers users access on the Internet.

This study showed that machine learning would effectively address an important gap between complex medical diagnostics and everyday awareness of personal health. ACHYUTA 2.0 provides early evidence of risk for diabetes, kidney disease, and COVID-19, giving users the opportunity to take the initiative in their health and seek further medical support at the appropriate time.

Future work will focus on:

- Adding to the disease prediction portfolio to include cardiovascular and respiratory conditions
- Real-time health monitoring with IoT devices
- Personalized recommendation systems generated from prediction outcomes
- Advancing model improvement through more advanced deep learning
- Clinical validation in the real-world context

References

[1] Esteva, A., et al. (2019). A guide to deep learning in healthcare. *Nature Medicine, 25*(1), 24–29.

[2] Deo, R. C. (2015). Machine learning in medicine. *Circulation, 132*(20), 1920–1930.

[3] Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *JAMA, 319*(13), 1317–1318.

[4] Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*(14), 1347–1358.

[5] Weng, S. F., et al. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One, 12*(4), e0174944.

[6] Mallol, R., et al. (2025). Metabolomic profile of severe COVID-19 and a signature predictive of progression towards severe disease status: A prospective cohort study. *Scientific Reports, 15*(1), 4963.

[7] Wu, I.-W., et al. (2025). Machine-learning assisted discovery unveils novel interplay between gut microbiota and host metabolic disturbance in diabetic kidney disease. *Gut Microbes, 17*(1), 2473506.

[8] Sarker, M. (2024). Revolutionizing healthcare: The role of machine learning in the health sector. *Journal of Artificial Intelligence General Science, 2*(1), 37–46.

[9] Kikuchi, K., et al. (2019). Gut microbiome-derived phenyl sulfate contributes to albuminuria in diabetic kidney disease. *Nature Communications, 10*(1), 1835.

[10] Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56.

[11] Wen, Y., Wan, Z., Ren, H., Wang, X., & Wang, W. (2025). Interpretable machine learning model for predicting and assessing the risk of diabetic nephropathy: Prediction model study. *JMIR Medical Informatics, 13*, e64979.

[12] Hsu, C.-T., Pai, K.-C., Chen, L.-C., Lin, S.-H., & Wu, M.-J. (2023). Machine learning models to predict the risk of rapidly progressive kidney disease and the need for nephrology referral in adult patients with type 2 diabetes. *International Journal of Environmental Research and Public Health, 20*(4), 3396.

[13] Li, J., Jin, Z., Zhan, Z., Huang, X., Sun, Y., Yin, S., Li, H., Hu, R., & Liu, Y. (2025). Machine learning-based risk predictive models for diabetic kidney disease in type 2 diabetes mellitus patients: A systematic review and meta-analysis. *Journal of Diabetes & Metabolic Disorders.*

[14] Arif, M. S., Rehman, A. U., & Asif, D. (2024). Explainable machine learning model for chronic kidney disease prediction. *Algorithms, 17*(10), 443.

[15] Bunn, A., Smith, R., & Lee, X. (2024). Predicting chronic kidney disease progression with artificial intelligence. *BMC Nephrology, 25*, 148.

[16] Hassanein, H., Radhakrishnan, Y., & Vachharajani, V. T. (2024). Machine learning for predicting chronic renal disease progression in COVID-19 patients with acute renal injury: A feasibility study. *Journal of Clinical Nephrology*, 48(2), 112–120

[17] Lei, N., Zhang, X., Wei, M., et al. (2022). Machine learning algorithms' accuracy in predicting kidney disease progression: A systematic review and meta-analysis. *BMC Medical Informatics and Decision Making, 22*, 205. https://doi.org/10.1186/s12911-022-01951-1 BioMedCentral

[18] Shivhare, A., Kumari, A. C., & Srinivas, K. (2025). Predictive modeling for chronic kidney disease: A comparative analysis of machine learning techniques with explainable AI for clinical transparency. *International Journal of Education and Management Engineering, 15*(3), 24–39. https://doi.org/10.5815/ijeme.2025.03.03 mecs-press.org

[19] Sivakumar, R., & Vijayalakshmi, R. (2024). A survey of predicting CKD using machine learning. *International Journal of Intelligent Systems and Applications in Engineering, 12*(4), 3176–3182. IJISAE

[20] Uddin, M. K., Akter, S., Das, P., & Anjum, N. (2024). Machine learning-based early detection of kidney disease: A comparative study of prediction models and performance evaluation. *International Journal of Medical Science and Public Health*, *5*(12), 58–75. https://doi.org/10.37547/ijmsphr/Volume05Issue12-05 ijmsphr.com

[21] Haque, M. E., Islam, S. M. J., Maliha, J., Sumon, M. S. H., Sharmin, R., & Rokoni, S. (2025). Improving chronic kidney disease detection efficiency: Fine-tuned CatBoost and nature-inspired algorithms with explainable AI. *arXiv*. https://doi.org/10.48550/arXiv.2504.04262 arXiv

[22] Iliyas, I. I., Saidu, I. R., Dauda, A. B., & Tasiu, S. (2020). Prediction of chronic kidney disease using deep neural network. *arXiv*. https://arxiv.org/abs/2012.12089 arXiv

[23] Hsu, C.-T., Pai, K.-C., Chen, L.-C., Lin, S.-H., & Wu, M.-J. (2023). Machine learning models to predict the risk of rapidly progressive kidney disease and the need for nephrology referral in adult patients with type 2 diabetes. *International Journal of Environmental Research and Public Health*, *20*(4), 3396. https://doi.org/10.3390/ijerph20043396 MDPI