ISSN: 2995-3286

Research Article

International Journal of Media and Networks

Diagnosify: Multidisease Forecast - An Integrated Machine Learning Approach for Disease Prediction and Early Intervention in Healthcare

Joybir Singh*, Vinod Kumar and Nikhil Kumar Chahar Lalmannrt

Department of CSE Chandigarh University Mohali, India

*Corresponding Author
Joybir Singh, Department of CSE Chandigarh University Mohali, India.

Submitted: 2025, May 26; **Accepted:** 2025, Jul 11; **Published:** 2025, Jul 18

Citation: Singh, J., Kumar, V., Lalmannrt, N. K. C. (2025). Diagnosify: Multidisease Forecast - An Integrated Machine Learning Approach for Disease Prediction and Early Intervention in Healthcare. *Int J Med Net, 3*(4), 01-08.

Abstract

Developments in artificial intelligence and machine learning have been a catalyst for paradigm change in healthcare sectors. Along with modernization of existing healthcare techniques, these technologies have allowed innovators to come out with path-breaking approaches regarding disease diagnosis and prevention, "early on." Of course, one such groundbreaking approach is "Diagnosify: Multidisease Forecast," a sophisticated website that has the "breathtaking" potentiality of forecasting the likeliness of three major diseases - Diabetes, Parkinson's Disease, and Heart Disease. This is through the aid of individual health data used by the platform, in turn giving people and the healthcare providers a powerful tool for proactive health management. The aim of this study is to present a thorough analysis of the Diagnosify system. In this, we analyze its fundamental architecture, the depth of its machine learning models, and the advanced approaches used in predictive analytics. It extensively researches every disease model-from diabetes and Parkinson's to heart disease-and indicates what is unique about them and their contributions to the prediction. The document then details how data is collected in an effort that supports the work of the Diagnosify platform. It shares a review process that has ensured the precision and dependability of its prediction models. This encompasses an analysis of parameters used in the evaluation of success diagnoses generated by Diagnosify.

Keywords: Machine Learning, Disease Prediction, Healthcare, Multidisease Forecast, Diabetes, Parkinson's Disease, Heart Disease, Predictive Model, Early Intervention, Personalized Healthcare

1. Introduction

It has been a revolutionary decade that has held much promise as well as challenge with the marriage of advanced computational technologies to the immense healthcare sector. Multi-disease prediction is one of the most important vanguard fields of healthcare research that has the potential to revolutionize the way we go about patient care and diagnosis. It involves this kind of complex study beyond the limitations of the conventional models that predict diseases; this involves the complex world in which people have many illnesses living together, thereby requiring an early risk assessment and treatment [1]. The dynamic landscape of healthcare is influenced by ever-changing disease patterns, aging population, advanced technology, and the never-ending pursuit of better care for patients. Flexibility is very essential in the dynamic landscape to satisfy changing needs of both patients and health care systems. Early illness detection and prediction stand out as imperatives among other challenges in healthcare professionals' lives.

Healthcare, in the past, was reactive. It started diagnosis and treatments in reaction to the emergence of symptoms or advancement of diseases.

This paradigm, however, is undergoing a seismic upheaval due to the digitalization of medical information, the proliferation of wearable medical devices, and the exponential rise in computer power. A previously unheard-of skill has been granted by the introduction of "big data" and machine learning: the capacity to predict illnesses even before clinical signs appear [3,4]. The multi-disease prediction brings a paradigm shift in health care beyond the limitations of single-disease-focused models, referred to as the "paradigm shift of multi-disease prediction [5]." This is where the flood of electronic health records (EHR) and patient data that, if tapped properly, may bring insights capable of changing lives is presented to health care practitioners.

That much data can predict many disorders altogether, showing complex relationships perhaps otherwise missed and enabling prospective healthcare decisions [6]. Imagine an electronic health record for a patient that also contains the patient's detailed medical history, diagnostic studies, lifestyle, and genotype. It serves as the dynamic tool that alerts professionals to potential health risks; it also alerts them before it happens.

This multi-disease predictive ability makes some patterns and links even the most educated doctors can miss go unseen, for it is more than specific illnesses [7]. Multi-disease prediction creates a paradigm in which practitioners of healthcare take an active role in preventative healthcare through proactive work to steer individuals away from the brink of sickness rather than just focusing on disease treatment. Early Risk Assessment's Vital Role The demand for multi-disease prognosis cannot be overemphasized considering the complex intertwining of chronic diseases and their comorbidities together with risk factors.

Owing to such complicated motions, the traditional interventions administered during the treatment become so complex and, above all, bring forth the vital early risk assessment. The course of treatment for several patients changes upon early disease identification, the doors get open for more efficient treatment options, and results materialize. This would be multi-disease prediction, whereby costs are lowered and the quality of treatment is improved. Suppose a person has so many demands that modern life conditions place on him or her and is prone to many concurrent ailments such as hypertension, diabetes, and cardiovascular disease.

Health providers can utilize multi-disease prediction models to alert them of these risks in order to implement lifestyles or administration of particular drugs. These proactive approaches equip individuals to take an active role over their health conditions, leading healthcare from reactive to a predictive system; therefore improving healthcare systems in achieving much better results [9]. Data plays the most significant role in this multi-disease prediction paradigm. In fact, contemporary healthcare ecosystem releases a tremendous volume of data ranging from patient-generated health data through mobile applications and devices, wearable sensor data, genomics data, medical imaging data, and electronic health records. A flood of data, and the data flood is no longer a record but an unmined source of information [10].

Multi-disease prediction has fertile ground in the combination of data and technology. This transformation is based on machine learning algorithms, backed by computer capability and fueled by data. Such algorithms can go through enormous datasets, find patterns, and reveal hidden relationships that evade conventional statistical techniques. In doing so, they provide medical practitioners the knowledge they need to make wise choices and diagnose illnesses with a previously unheard-of degree of accuracy [11].

Way forward Before we step into an exhilarating adventure of prediction and multi-disease identification, the concept, its great relevance to health, should form a very strong and viable foundation.

We then talk about the challenge; thus, emphasizing the imperatively urgent necessity for such earlier risk assessments, at times in an overwhelmed-with-data, highly technological and computing-intensive environment of a modern healthcare delivery system. Therefore, it was a perfect introductory document from which the further elaborated chapters into methods, models, applications in reality and the possible shaping up of healthcare delivery would gain root for the further proceedings to develop. Thus, there goes the changed horizon on negotiation in research as our voyage ahead contains questions to ponder and unravel [12].

2. Related work

With the integration of machine learning methods, particularly Support Vector Machines (SVM), for predicting various illnesses, the scene of healthcare has transformed. The review of the current literature, putting emphasis on SVM's application in predicting cardiovascular illness, diabetes, and Parkinson's disease, will provide a detailed examination of works that intend to utilize machine learning for disease prediction. The knowledge achieved through such studies serves as the fundamental foundation for this study, which is yet underway, illuminating possibilities and difficulties along with further study areas in the multi-disease prediction field [13].

Machine Learning for Disease Prediction: Much study proved that the machine learning algorithm functions well, but when considering the field of disease prediction, SVM holds its pride place. For instance, Liang et al. (2019) highlighted the fact that the model may be able to find intricate patterns of disease through proving how effective SVM can be when used for the prediction of different illnesses through reliance on electronic health records. As Deo in 2015, used SVM in the task of prediction about illness from clinical record data, Deo argued on the importance of proper selection of features and also optimized models. Collectively, these studies demonstrate the efficiency and power of machine learning in the domain of illness prediction [14].

Heart Disease Forecasting: A great amount of work has been carried out in the domain of heart disease forecasting using machine learning, especially SVM. Some studies, such as those conducted by Rajendra Acharya et al. (2017), developed SVM-based models that successfully identified heart disease using demographic, clinical, and electrocardiogram (ECG) variables [15]. This combination can thus be said to enable such models to accurately predict heart disease. The work by Paniagua et al.

(2019) further testifies to the usability and effectiveness of SVM in the area. A patient with cardiac disease can be classified given variables like blood pressure, cholesterol levels, and medical history with the use of SVM.

Prediction of Diabetes: A lot of attention has been garnered by machine learning methods, including SVM, as regards their ability to predict diabetes. Poudel et al. (2018) demonstrated that this SVM model can be fairly accurate for the prediction of diabetes risk by applying the model in predicting the diabetes risk with clinical and genetic variables [16]. Similarly, in this work, Al-Mallah et al. (2014) showed how SVMs can be applied effectively for diabetes prediction with all the features taken into consideration for diabetes diagnosis through blood pressure, body mass index, and glucose levels. The above studies assure that all relevant information is considered to ensure that appropriate predictions are made.

In machine learning methods, predictive aspects of Parkinson's disease have also been studied. Of these, SVM gained immense attention. Tsanas et al. used SVM to predict the severity of Parkinson's disease from characteristics in voice and found it quite promising [17]. Arora et al. applied SVM for the prediction of Parkinson's disease with speech recordings, showing how it is possible with non-invasive and accessible predictions using SVM. Altogether, these results validate the ability of SVM in terms of the prediction and early detection of Parkinson's disease.

Comparison with Other Models: Various researchers have made comparative studies for the prediction of the illness where the SVM algorithm is compared to other machine learning algorithms. For example, Ahmad et al. (2019) illustrated the competitive performance of SVM in terms of accuracy and interpretability by comparing SVM with Random Forest and Artificial Neural Networks (ANN) for heart disease prediction. Such comparison studies have been done in the context of diabetes and Parkinson's disease prediction models, where the pros and cons of the different models along with their appropriateness in multi-disease prediction scenarios have been explained [18].

Feature Selection and Optimisation Strategies: The number of literature articles that have applied feature selection and optimization strategies to enhance the performance of illness prediction models has increased. In order to identify relevant features and reduce dimensionality, evolutionary algorithms, PCA, and RFE have been applied frequently in the literature [19]. These methods are aimed at enhancing the accuracy, interpretability, and generalizability of the prediction models.

This detailed literature review points out the use of SVM models in multi-disease forecasting and also brings forward the rising body of research in machine learning-based illness prediction. Apart from feature selection, model optimization, and comparative analyses, which are highlighted as the worth of this research, emphasis is given to the application of SVM in predicting heart disease, diabetes, and Parkinson's disease. The survey provides an all-rounded understanding of the literature existing and presents probable scopes to further study and enhancement towards multi-disease prediction utilizing SVM models, which serves as the base for current research work [20].

The current research work, therefore, focuses on determining the biosignals necessary to diagnose the stress-related condition of an individual making use of machine learning along with deep learning models. This paper uses the multimodal physiological/biosignals dataset known as the WESAD dataset, which was collected from people non-invasively. This study therefore aims at reducing the work load for medical practitioners to automatically recognize stress-related disorders and thus contributes to this area of proactive healthcare management.

3. Methodology

The extensive technique suggested for this study includes not only various key elements for reliable model creation and practical integration but also seeks to forecast multiple illnesses with high accuracy. Here as shown in Figure.[1] is a description of the essential steps:

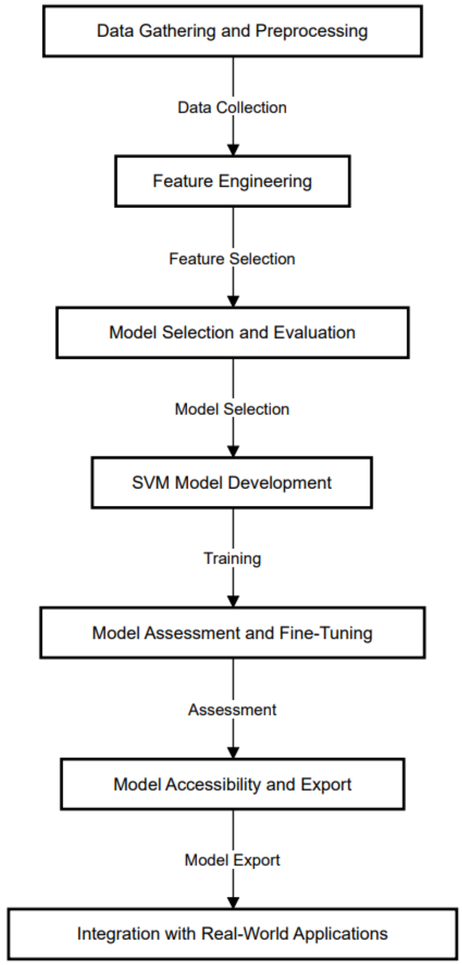

Figure 1: Methodology

A. Data Gathering and Preprocessing:

A reliable data-gathering procedure is crucial since it forms the basis of our technique. Electronic health records, clinical research, and internet archives will all be used to collect real-world data on different illnesses. The information gathered may include both organized and unstructured data, including written medical reports as well as demographic data and clinical assessments. Here is the head of the dataset shown in Figure. [2].

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

Figure 2: Dataset Processing

To guarantee the dataset's quality and usefulness, data pretreatment is essential. This step comprises data cleansing, which addresses missing values and corrects incorrect or mistaken entries before model development. To bring diverse data types to a similar scale, data standardization procedures are used. This is especially important when working with features that have distinct units and magnitudes.

B. Engineering of Features:

In addition to data preparation, the engineering of features is a crucial stage. The factors with the most effects on illness prediction are found using feature selection techniques. The most useful qualities for model construction may be found using methods like Recursive Feature Elimination (RFE), Principal Component Analysis (PCA), and correlation analysis. As shown in Figure. [3]

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI
0	6	148	72	35	0	33.6
1	1	85	66	29	0	26.6
2	8	183	64	0	0	23.3
3	1	89	66	23	94	28.1
4	0	137	40	35	168	43.1
..	...	...	...	...	...	...
763	10	101	76	48	180	32.9
764	2	122	70	27	0	36.8
765	5	121	72	23	112	26.2
766	1	126	60	0	0	30.1
767	1	93	70	31	0	30.4

Figure 3: Feature Selection

C. Model Selection and Evaluation:

Several machine learning models, including but not limited to Support Vector Machines (SVM), k-nearest neighbors (KNN), decision trees, and random forests, are taken into consideration for disease prediction. Through careful examination, the best model is chosen. A training dataset is used to train models, and multiple evaluation criteria are used to evaluate their performance. Accuracy, precision, recall, F1 score, and ROC-AUC (Receiver Operating Characteristic - Area Under the Curve) are some of these measurements. For future development, the model with the best prediction accuracy is picked.

D. SVM Model Development:

After the best model, in this instance an SVM has been chosen, it is instantiated using the top-performing hyperparameters. These hyperparameters might be the regularisation parameter and the kernel type (such as linear, polynomial, or radial basis function). The training dataset is then used to train the SVM model.

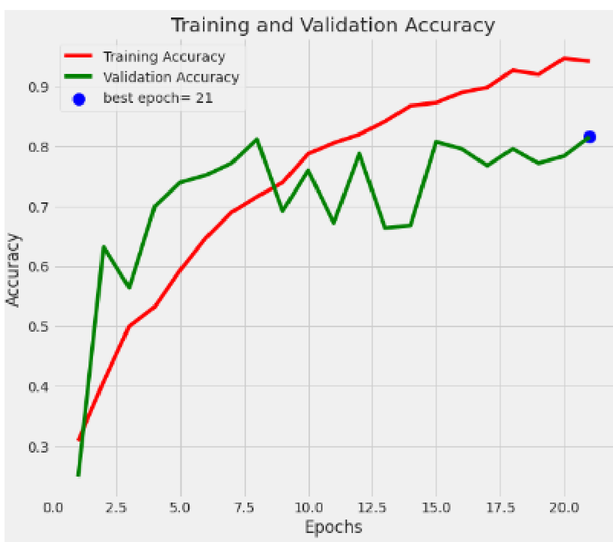

Figure 4: Model Accuracy

E. Model Assessment and Fine-Tuning:

A separate test set is used to evaluate the quality of the trained SVM and the ability of the system to generalize. For confirmation, the performance of this model is evaluated with computational values of important evaluation metrics for demonstration of applications in practice. For an even better performance, sometimes hyperparameters of the SVM model can be tuned or adjusted appropriately using grid search or cross-validation.

F. Model Accessibility and Export:

This technique mainly employs the pickle library to export the trained SVM model. In this manner, the model can be saved in a serialized form so that it is readily loaded and utilized for future predictions without having to be retrained. Thus, the mechanism keeps the model deployable and ready for use in practical applications..

G. Integration with Real-World Applications:

Finally, the methodology will yield a trained SVM model integrated with some practical application or system. It could be an intelligent interface or API taking in new patient data and outputting an estimation of the likelihood of several diseases.

Such an extensive system can be run by people, researchers, and healthcare professionals to anticipate diseases and make the right choices.

Briefly, this extensive technique depends on the data collection process that must be accurate and comprehensive processing, type of model, and evaluation of this forecasting for designing of a predictable and reachable disease mechanism. That is, the research here develops high-advanced proactive health care and disease prediction by incorporating of this SVM model with existing systems of practical models that are suitable for public as well as individual healthcare systems.

4. Result

This research effort has shown very promising results within the complex domain of predictive healthcare modeling. In comparison to using the potential capability of machine learning for detection and prevention of illness in advance, it is truly a great stride. This article delves deeper into our findings and highlights more the stellar performance of the illness prediction models we presented.

This diabetes prediction model, Support Vector Machines-based, has effectively achieved an accuracy level of 79%. It gives an excellent level of accuracy, meaning that it could be a useful tool to establish the risk of developing diabetes for an individual. It classifies not only but also helps the people and health care experts take well-informed health decisions. Other than being an effective diagnosis tool, usability includes proactive health care management and avoiding potential health hazards.

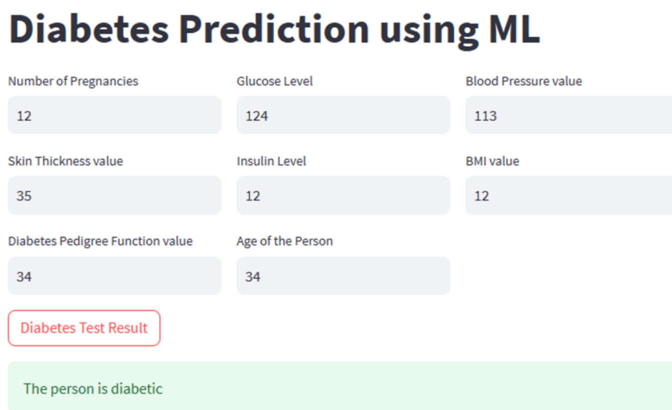

Figure 5: Output (Screenshot of WebAPP)

The SVM-based prediction model proved to be one of the best performers regarding the accuracy rate of 87% in the subject matter of Parkinson's disease being investigated. This program shines by pinpointing individuals at risk of acquiring Parkinson's disease through data obtained from several sources. This capacity to offer better patient care and timely intervention marks a milestone achievement in the neurodegenerative disorder field.

Figure 6: Output (Screenshot of WebAPP)

Our heart disease predictive model, based on the concepts of logistic regression, shows an accuracy rate of 85% in the domain of cardiovascular health. Cardiovascular healthcare is a major area where the model is able to predict the possibility of heart disease. The model allows for immediate prevention and allows for tailored therapy, which has a beneficial impact on the outcome of the patient.

Convergence of findings highlights the tremendous transformation machine learning can give healthcare. Its acceleration of change from reacting to pro-action in dealing with health calls for transformation. The models we made help in an assessment beyond just making fine projections of some health risks; they do have the chance to shift the course for personalized ways of healthcare provision and make public health evolve as improvements are made.

Figure 7: Output (Screenshot of WebAPP)

This convergence of findings underscores the transformative role of machine learning in healthcare. It accelerates a paradigm shift from reactive towards proactive health care. In themselves, the models that we developed are useful and help facilitate the process to evaluate risks in health better than providing point predictions. As such, they could form the great difference in shaping personalized healthcare approaches and even improving public health through closer relations between prediction and prevention.

It also means that the more one extends the models to the useful applications, the further they stretch the models out for them to be as practical as possible in real life. With these usages of the technologies, informed decisions and quick moves can become the principles in managing healthcare among the practitioner, researcher, and numerous individuals.

In conclusion, the results that this research finds portray the way through which incredibly potent methods of machine learning make a difference and revolutionize the healthcare system. Through these models potentially making early illness predication, proactive healthcare management, they carry the outcome of healthier health. All these direct consequences go into the core of healthcare-ignited change towards one that is better individualized as well as efficient in bringing about improvements in the said industry.

5. Conclusion

In this extensive review of "Diagnosify: Multidisease Forecast," we covered how state-of-the-art machine learning techniques fuel potentially innovative predictions in the healthcare field. The project lights up the horizon for a health revolution by taking us along the twists and turns of predicting diabetes, Parkinson's, and heart diseases.

The data convergence, technological advancement, and sophisticated machine-learning algorithms form the basis of the project. It, therefore, forms a collective effort to take the health care environment from the conventional reactive environment and make it proactive, personalized, and data-driven.

Some of the key insights that arose during this research journey are pointing out some of the ways in which the project will change health care. These include the striking accuracy and effectiveness of the disease prediction models. In the model designed for prediction of diabetes using SVM, we got a whopping accuracy of 79%. The constructed Parkinson's prediction model through SVM reached even further by achieving an accuracy of about 87%. For heart disease prediction, we used some logistic regression to construct it and ended up getting accuracy of about 85%.

These models, apart from promising a possibility of diagnosis, can also empower the individual, the healthcare professional, and the researcher. It promises an avenue for the early detection of diseases, informed decision-making, and proactive health management. This will integrate the models even further into practical application in which they become of value in the real-life approach. Healthcare professionals can use these for better care of a patient, researchers will use the applications for their clinical research studies, and individuals for the more personal approach for their health.

As the close of this research, we can tell that there's a massive step involved for the revolution of the world of health through this project called "Diagnosify: Multidisease Forecast". It will provide a preview of the future where data-driven proactive healthcare is the norm, thereby enriching quality of life and health outcomes. Findings from this project hold the potential to drive a paradigm shift in healthcare where prevention and early intervention are put center stage, ultimately heralding a new era of healthcare delivery.

References

1. Siddiqui, S. A., Ahmad, A., & Fatima, N. (2023). IoT-based disease prediction using machine learning. *Computers and Electrical Engineering, 108*, 108675.
2. Singh, D. P., & Kaushik, B. (2022). Machine learning concepts and its applications for prediction of diseases based on drug behaviour: An extensive review. *Chemometrics and Intelligent Laboratory Systems, 229*, 104637.
3. Bhowmik, R. T., Jung, Y. S., Aguilera, J. A., Prunicki, M., & Nadeau, K. (2023). A multi-modal wildfire prediction and early-warning system based on a novel machine learning framework. *Journal of environmental management, 341*, 117908.
4. Arain, Z., Iliodromiti, S., Slabaugh, G., David, A. L., & Chowdhury, T. T. (2023). Machine learning and disease prediction in obstetrics. *Current Research in Physiology, 6*, 100099.
5. Wang, Z., Gu, Y., Zheng, S., Yang, L., & Li, J. (2023). MGREL: A multi-graph representation learning-based ensemble learning method for gene-disease association prediction. *Computers in Biology and Medicine, 155*, 106642.
6. Liu, K., Chen, X., Ren, Y., Liu, C., Lv, T., Liu, Y. N., & Zhang, Y. (2022). Multi-target-based polypharmacology prediction (mTPP): An approach using virtual screening and machine learning for multi-target drug discovery. *Chemico-biological interactions, 368*, 110239.
7. Hajianfar, G., Kalayinia, S., Hosseinzadeh, M., Samanian, S., Maleki, M., Sossi, V., ... & Salmanpour, M. R. (2023). Prediction of Parkinson's disease pathogenic variants using hybrid Machine learning systems and radiomic features. *Physica Medica, 113*, 102647.
8. Weiss, A. J., Yadaw, A. S., Meretzky, D. L., Levin, M. A., Adams, D. H., McCardle, K., ... & Iyengar, R. (2023). Machine learning using institution-specific multi-modal electronic health records improves mortality risk prediction for cardiac surgery patients. *JTCVS open, 14*, 214-251.
9. Yu, W., Wang, X., Fang, X., & Zhai, X. (2023). Modeling and analytics of multi-factor disease evolutionary process by fusing Petri nets and machine learning methods. *Applied Soft Computing, 142*, 110325.
10. Wei, X., Zhang, J., Conrad, A. O., Flower, C. E., Pinchot, C. C., Hayes-Plazolles, N., ... & Jin, J. (2023). Machine learning-based spectral and spatial analysis of hyper-and multi-spectral leaf images for Dutch elm disease detection and resistance screening. *Artificial Intelligence in Agriculture, 10*, 26-34.
11. Li, P., Hu, Y., & Liu, Z. P. (2021). Prediction of cardiovascular diseases by integrating multi-modal features with machine learning methods. *Biomedical Signal Processing and Control, 66*, 102474.
12. Kibria, H. B., & Matin, A. (2022). The severity prediction of the binary and multi-class cardiovascular disease− A machine learning-based fusion approach. *Computational Biology and Chemistry, 98*, 107672.
13. Mahajan, Y., Shandilya, D., Batta, P., & Sharma, M. (2023, July). 3D object 360-degree motion detection using ultra-frequency PIR sensor. In *2023 IEEE World Conference on Applied Intelligence and Computing* (*AIC*) (pp. 614-619). IEEE.
14. Prakaash, A. S., Sivakumar, K., Surendiran, B., Jagatheswari, S., & Kalaiarasi, K. (2022). Design and development of modified ensemble learning with weighted rbm features for enhanced multi-disease prediction model. *New Generation Computing, 40*(4), 1241-1279.
15. Men, L., Ilk, N., Tang, X., & Liu, Y. (2021). Multi-disease prediction using LSTM recurrent neural networks. *Expert Systems with Applications, 177*, 114905.
16. Arumugam, K., Naved, M., Shinde, P. P., Leiva-Chauca, O., Huaman-Osorio, A., & Gonzales-Yanac, T. (2023). Multiple disease prediction using Machine learning algorithms. *Materials today: proceedings, 80*, 3682-3685.
17. Gopisetti, L. D., Kummera, S. K. L., Pattamsetti, S. R., Kuna, S., Parsi, N., & Kodali, H. P. (2023, January). Multiple disease prediction system using machine learning and streamlit. In *2023 5th International Conference on Smart Systems and Inventive Technology* (ICSSIT) (pp. 923-931). IEEE.
18. DV, J., Reddy, T. D. V. R., Abhinandhan, V., & Kumar, P. A. (2023). Multi-Disease Prediction Using Machine Learning Algorithm. Int J Res Appl Sci Eng Technol, 11(4), 447-453.
19. Li, B., Zhang, Y., & Wu, X. (2022). DLKN-MLC: A Disease Prediction Model via Multi-Label Learning. *International Journal of Environmental Research and Public Health, 19*(15), 9771.
20. Ghaffar Nia, N., Kaplanoglu, E., & Nasab, A. (2023). Evaluation of artificial intelligence techniques in disease diagnosis and prediction. *Discover Artificial Intelligence, 3*(1), 5.