

An effective Leveraging Ensembling Methods for High-Enactment Chronic Disease Prediction Systems

Manuel R. Tanpoco¹, Preeti Patil², Dr Uttara Patnaik³, Aniruddha Bodhankar⁴, Senthil Kumar A⁵,
T R Vijaya lakshmi⁶

Submitted: 22/04/2023

Revised: 18/06/2023

Accepted: 28/06/2023

Abstract: A large portion of the Indian population cannot simply access healthcare facilities. Private hospitals located in and around towns and cities are the primary providers of medical care. Because of this, patients in remote towns and villages must travel great distances to receive basic and specialized medical care. It has been noted that consumers only seek medical attention from a doctor or other medical professional when all other home cures have failed to relieve their symptoms. This is likely due to a lack of understanding as well as other behavioural variables. A novel approach to ensembling was created by the researchers. Results from experiments validate the effectiveness of the proposed strategy in boosting the classification accuracy. The disease prediction system developed in this study showed promising results and has use in the early detection of chronic diseases. The increasing death toll from chronic illness can be combated with the help of these prototypes. The introduction of these chronic disease prediction systems in primary care settings would be a major improvement in healthcare quality. The scope of this study will be broadened to incorporate the creation of prediction models for other diseases.

Keywords: SVM; LR; CVD; Chronic diseases.

1. Introduction

The main causes of death in India are chronic illnesses like cancer, diabetes, and CVDs. The World Health Organization (WHO) urges the urgent development of affordable drugs for the treatment and prevention of various illnesses. The second leading cause of death worldwide is cancer [1]. Among a population of 13 billion people, slightly more than ten lakh new instances of cancer are identified each year. In addition to cancer, a number of other chronic diseases represent a serious threat to the health of the Indian population. Concerns are raised about the high frequency, morbidity, and mortality of chronic kidney disease (CKD). A global health issue is the viral illness hepatitis. In India, 3.3% of people are infected with hepatitis B. Over 40 million

people in India have hepatitis [2]. The national prevalence of diabetes was found to be 11.8 percent in 2019 according to the National Diabetes and Diabetic Retinopathy Study. The World Health report points out that among the major NCDs, the increase in diabetes-related health losses in India since 1990 has been the greatest. All of these statistical data indicate that there is room for improvement in the Indian healthcare industry. Healthcare facilities that are accessible and affordable will help the problem [3].

Since information technology (IT) has been fully incorporated into the healthcare industry, it has seen greater evolution. The goal of integrating IT into healthcare is to improve people's quality of life by making it more comfortable and inexpensive, similar to how cell phones did so. Making healthcare smarter could make this possible. For instance, the development of smart ambulances, hospitals, and other healthcare infrastructure could do this [4]. Every year, a study on patients affected by chronic diseases in a certain area is conducted, and it is discovered that there is very little gender difference in the patients. It is also discovered that many patients were admitted in 2014 in order to treat their chronic illnesses. Instead of using simply structured data, using both structured and unstructured data yields findings that are extremely accurate [5-6].

Recent decades have seen the introduction of some groundbreaking technologies for the quick collection of

¹Assistant Professor, Department of Decision Sciences and Innovation, De La Salle University, Email Id: manuel.tanpoco@dlsu.edu.ph Orcid Id: 0000-0002-0009-8250

²Associate Professor, Department of Information Technology, D.Y.Patil College of Engineering, Akurdi, Pune, India. Email ID: dr.preetipatil.dypa@gmail.com Orcid Id: 0000-0001-6304-4551

³District Program Manager, National Ayush Mission, Orissa, India, Email Id: patnaikdruttara@gmail.com

⁴Department of Management, Assistant Professor, Dr.Ambedkar Institute of Management Studies & Research,Nagpur, Email Id : aniruddha.bodhankar16@gmail.com, Orcid id : 0000-0001-8721-5714

⁵Professor, Artificial intelligence, University and city with country name: Sri Vishnu Engineering College for Women, Email Id: drsenthilkumar.ai@svecw.edu.in Orcid Id: 0000-0003-2567-7950

⁶Associate Professor, ECE, Mahatma Gandhi Institute of Technology, Email: trvijayalakshmi_ece@mgit.ac.in Orcid Id : 0000-0002-1107-2025

data, including MRI (magnetic resonant imaging) readouts, ultrasonography, social media data, and electronically acquired activity, behavioural, and clinical data. Because these healthcare big data sets are high dimensional, the number of features captured for each observation may exceed the number of observations as a whole [7]. They lack statistical power, are cross-sectional, noisy, and scarce. The problems in the high-dimensional data sets can be resolved by using machine-learning techniques. In a number of fields, machine learning contributes more. At the same time as healthcare epidemiology is on the verge of a significant transition, many complex models employ greater training data that are already available [8]. This data can help increase understanding of disease risk factors, decrease healthcare-associated infections, improve patient risk stratification, and identify infectious disease transmission routes. The interpretation of laboratory findings and other patient information can be made easier by machine learning to aid in the early detection of diseases [9]. Using knowledge discovery in the database, the low-level data might be transformed into high-level knowledge to help with early disease identification. To improve prediction accuracy and cut down on model training time, the data collected to create a data set should be pre-processed for missing values before only the crucial characteristics required for precise illness prediction are selected [10].

Diabetes, cardiovascular conditions, cancer, strokes, hepatitis C, and arthritis are the most prevalent chronic illnesses. The detection of such diseases is crucial in the healthcare industry since they are long-lasting and have a high death rate. Predicting the condition can help people take preventative measures to avoid having it, and early discovery of the disease can aid with improved treatment [11]. Machine learning uses a variety of methods, including supervised, semi-supervised, unsupervised, reinforcement learning, evolutionary learning, and deep learning. The processing of features that have been retrieved from real data and are organized as vectors is the issue at hand. The effective fusion of such vectors determines the processing quality. However, the high dimensionality of the vectors or the inconsistencies in the data are typically major problems [12]. In order to transform the data set into a highly compatible dimension, it is crucial to minimize the dimensionality of the data set, even if doing so results in a slight loss of features. The performance of the model is enhanced by this decrease in the dimensionality of the data set [13].

What follows is the outline for the rest of the paper. The related work is briefly described in part 2, and the methodology and the theoretical foundations of the methods used are described in section 3. The simulation results and analysis are presented in section 4. For the

chapter's final section, "key findings" we summarize the most important results.

2. Previous Related Work

It is also crucial to apply feature extraction techniques to improve the system's performance. The process of extracting features involves transforming the original data into a data set with fewer variables that contain the most discriminating information. Medical datasets typically have many features and are high-dimensional. The processing of high-dimensional medical data sets necessitates a significant amount of computing power [14]. Furthermore, it takes a long time to process these datasets. The method of feature extraction reduces the number of features required for processing without sacrificing crucial or pertinent details for a particular disease. The redundant characteristics that are present in the disease datasets are also reduced via feature extraction. By using specific transformation techniques, it converts the data from high-dimensional to low-dimensional space [15]. Combining a portion of new features with existing features and generating new features from the original features to improve classification accuracy, decreases the dimensionality of the data. Also, data reduction speeds up learning and improves the effectiveness of any machine learning process. A crucial step in creating ML models for medical datasets is feature extraction. These datasets are typically very large and take a long time to process because of their high dimensionality [16].

A small group of academics conducted research to identify important risk factors for ischemic heart disease. The authors used the PCA method to find relationships between the attributes. Another group of researchers used the PCA method in their study to improve the efficacy of machine learning methods. A few additional researchers used PCA and fuzzy approaches to produce the unique feature extraction technique known as Adaptive Neuro-Fuzzy Inference System (ANFIS) [17]. The purpose of the investigation was to identify lymph disorders. The study's classification accuracy was 88.83%, and the results were satisfactory. Several researchers looked at the efficiency of the PCA-SVM combo for detecting coronary artery disease early (CAD). The research found that the PCA-SVM method improves accuracy [18]. Some scientists utilizing PCA offered an adaptive analysis of the data. Another team of researchers used the least square support vector machine classifier in conjunction with PCA to create a high-performance hepatitis detection system. The selection of hybrid features has been done using a variety of techniques [19]. The research carried out by a few more academics suggested a hybrid paradigm for the early detection of breast cancer. In order to determine the

significant features, the study found that a mix of sequential backward search and PCA techniques generated the best results. With the use of a hybrid feature selection and neural network, the study achieved an accuracy of 97.5% [20].

Recently, a group of researchers used the ReliefF- SVM-based approach for early breast carcinoma diagnosis. A 90% accuracy rate was attained by the diagnostic system. The system's sensitivity and specificity results were both 98.7%. The study's findings indicated that performance would improve with less dimensionality. Another group of researchers suggested a hybrid model called PCA-KNN that uses PCA and the K-Nearest Neighbor classifier [21]. The prediction algorithm created in this study had an astounding 88.3% accuracy in identifying liver disorders. To find the best potential subsets, F-score feature selection, and the Sequential Forward Search approach were used. Grid search was used to find the kernel's optimal hyperparameters in order to enhance the SVM classifier's performance [22]. A group's study suggested using PCA and SVM classifiers to differentiate between CAD and normal circumstances. Many researchers have proposed a brand-new method for diagnosing breast cancer. The authors proposed a method known as "RS SVM," or Rough Set Feature Selection and Support Vector Machine. The research has a 93% categorization accuracy rate [23].

Researchers have conducted a variety of studies to create machine learning-based prediction systems for cardiac disorders. The prediction algorithm created for this study revealed information that was concealed in the cardiac dataset. The method could be used to diagnose heart ailments early on. With a performance of 79.3%, Naive Bayes was shown to perform the best [24]. In a study conducted by a different team, a CVD prediction model was created utilizing a fuzzy strategy based on membership function. The accuracy, precision, and recall of the performance were assessed. The outcomes of the experiment demonstrated that the fuzzy k-NN classifier performed better than k-NN. A heart disease prediction system was built by a few more researchers employing 13 important criteria. The ANN method was employed to forecast CVDs. The accuracy of the suggested study was 80%. In their study, several researchers suggested an automated method for resolving challenging issues related to the diagnosis of heart disease. 80.6% accuracy was reached [25].

3. Purpose of the work

1) Examining how feature selection and ensembling methods can be used to improve the effectiveness of chronic disease prediction systems.

4. The Projected Work:

In order to improve the accuracy of prediction algorithms used in the early detection of chronic diseases like cardiovascular disease, diabetes, and chronic kidney disease, a new ensembling approach has been developed in this study. Logistic regression, naive Bayes, support vector machine, and decision tree are the foundational machine learning classifiers utilised here. All models have been assessed for their accuracy, sensitivity, specificity, positive predictive value, and negative predictive value. Prediction model performance was shown to be vastly enhanced by employing unique ensembling techniques and layered generalisation. Millions of lives could be saved with the help of these high-performance ensemble-based machine learning models if they were implemented in primary healthcare centres.

Here is a detailed explanation of these data sets:

1. The Cleveland heart disease dataset was used to investigate cardiovascular diseases. The purpose of the test is to identify cardiac conditions in patients. Predicting the occurrence of heart disease in this dataset requires input from 13 different variables. There are 303 records in the collection.

2. This work contributes to the UCI Pima Diabetes Dataset Repository. The presence or absence of diabetes in a given patient can be predicted using this dataset's 8 input independent medical variables. This dataset has 768 records.

3. Studies on the UCI chronic renal disease dataset have been conducted. There are 400 entries in this data set. In order to determine whether or not a patient is infected, we use 24 input attributes.

Let the dataset be signified as

$$D = \{A_j + B_j, \quad \text{where,} \quad j=1,2,3 \dots M\} \quad (1)$$

Input characteristics. Depending on how you choose to express the independent variables,

$$A_j = \{A_{j1}, A_{j2}, \dots, A_{jp}\} \quad (2)$$

The presence (1) or absence (0) of sickness is represented by the project's output attribute. It could also be spelled B_j . As a result, it appears that:

$$B_j = \{0, 1\} \quad (3)$$

Take into account a single distinguishing feature of multiple rows or records.

$$A_i = A_1 \dots A_n \quad (4)$$

One such representation of the weighted mean is:

$$A_t = \frac{1}{N} \sum_{k=1}^N w_k A_k \quad (5)$$

Where, $w_k \geq 0$

$$\sum_{k=1}^N W_k = 1 \quad (6)$$

A data factor with a higher weightage has a greater impact on the weighted mean than one with a lower weightage. We can only use weights with positive values. Weights cannot be negative, although some values may be zero, as seen in the preceding equation. Values are generated at random from the set [1, 2, 3, ..., p] in the proposed process, ensuring randomness. There is less uncertainty and more precision when results are generated randomly. Using an average-based partitioning rule, the dataset DS is split in half.

$$D = \begin{cases} \{D_{11} & \text{if } A_k < A_t \\ \{D_{12} & \text{if } A_k \geq A_t \end{cases} \quad (7)$$

Each child node must be treated separately. The nodes D11 and D12 are the primary ones. D11 gives rise to D21 and D22, while D12 gives rise to D23 and D24, as can be shown by applying the aforementioned partitioning method.

It is possible to continue data partitioning forever until some stopping criterion is reached. The two cut-offs used in this strategy are as follows:

1) Assume that Dmax is the greatest depth of the tree that can be used. When the depth of the tree reaches Dmax, like in the first halting criterion, further growth is halted. The root node has a depth of 0, D11 and D12 have a depth of 1, and so on.

2) If there are fewer instances in a partition than a specified threshold, growth of the tree is halted. Take into account that Nmin is the cutoff for defining the smallest feasible data partition size. If the number of children of a node, Dmn, is fewer than Nmin, then the tree will cease expanding. Thus, the criteria for terminating can be summed up as

$$D \geq D_{max} \text{ or } N(D_{mn} < N_{min}) \quad (8)$$

The Gini impurity index is found by plugging in the total number of classes, K, and the probability, p(m), of selecting a data item from class m.

$$G = \sum_m^k P(m) * (1 - P(m)) \quad (9)$$

The Accuracy-Based Weighted Age-Classification Ensemble (AB-WAE) has been used to create a set of ensemble classifiers.

$$W = \begin{cases} P_a & \text{if } P_a > P_a^{\pi} \\ \frac{P_a}{I} & \text{elsewhere} \end{cases} \quad (10)$$

The number I indicates how many times I was a part of the ensemble. Where Pa is the ensemble-wide median classification accuracy.

This ensemble method gives each classifier a varying weight based on its performance and duration in the ensemble. If a classifier's weight w falls below a certain value, it is removed from the ensemble. Since accuracy is employed as the optimisation criterion, the ensemble is able to produce the best possible results.

4. Result and Discussion:

Validating the model's efficacy is essential. Here, we employed several indicators to verify the system's efficacy.

5.1. Accuracy:

It is a typical metric for classifying test results numerically. Increased precision indicates a more efficient system.

$$\text{Accuracy} = \frac{TN+TP}{\text{Total data Sample}} \times 100 \quad (11)$$

5.2. Specificity:

Specificity was defined as the absence of incorrect data classification. True Negative Rate is another name for it (TNR). Figure IV displays the recall of the current method in comparison to commonly utilized methods.

$$\text{Specificity} = \frac{TN}{TN+FP} \times 100 \quad (12)$$

5.3. Sensitivity:

The accuracy with which the model places the test data into one of its classes constitutes the present method's sensitivity. How many true positives were successfully detected was the question it addressed. True Positive Rate is another name for it.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \times 100 \quad (13)$$

The cases labelled TP and TN were correctly predicted, while those labelled FN and FP were not. The table clearly shows that TP+FN equals the overall number of heart patients, whereas TP+FP stands for the expected number of people with heart disease.

Table 1: Accuracy Performance Comparison of Different classifiers with the projected method.

Classifiers	Accuracy (%)		
	Cardiovascular diseases	Diabetes	UCI chronic disease
Logistic Regression	79.2	89.6	82.4
SVM	85.6	78.9	80.7
Naïve Bayes	81.4	85.3	82.7
Decision Tree	81.8	84.6	82.1
Proposed Method	87.8	92.4	86.4

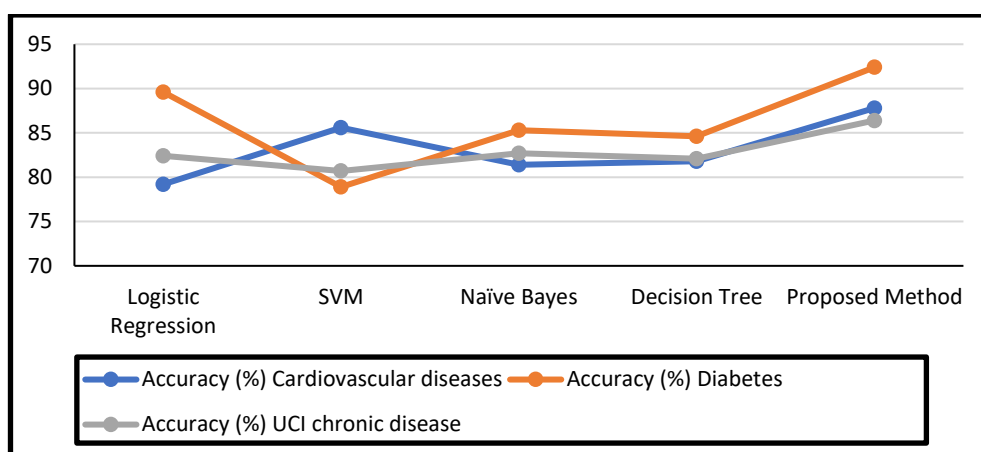


Fig 1: Accuracy Performance Comparison of Different classifiers with the projected method.

Table2: Sensitivity Performance Comparison of Different classifiers with the projected method.

Classifiers	Sensitivity (%)		
	Cardiovascular diseases	Diabetes	UCI chronic disease
Logistic Regression	78.2	83.4	81.6
SVM	87.5	78.6	81.6
Naïve Bayes	89.3	86.7	81.2
Decision Tree	82.7	85.3	80.9
Proposed Method	88.5	90.7	89.4

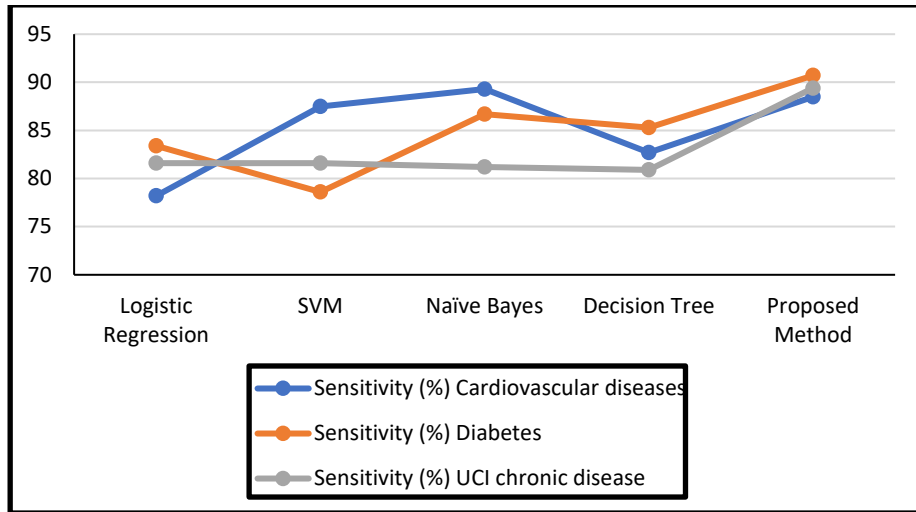


Fig 2: Sensitivity Performance Comparison of Different classifiers with the projected method.

Table 3: Specificity Performance Comparison of Different classifiers with the projected method.

Classifiers	Specificity (%)		
	Cardiovascular diseases	Diabetes	UCI chronic disease
Logistic Regression	85.4	88.3	80.7
SVM	89.2	80.6	80.7
Naïve Bayes	81.5	82.7	89.4
Decision Tree	81.4	80.6	89.8
Proposed Method	89.3	85.7	90.4

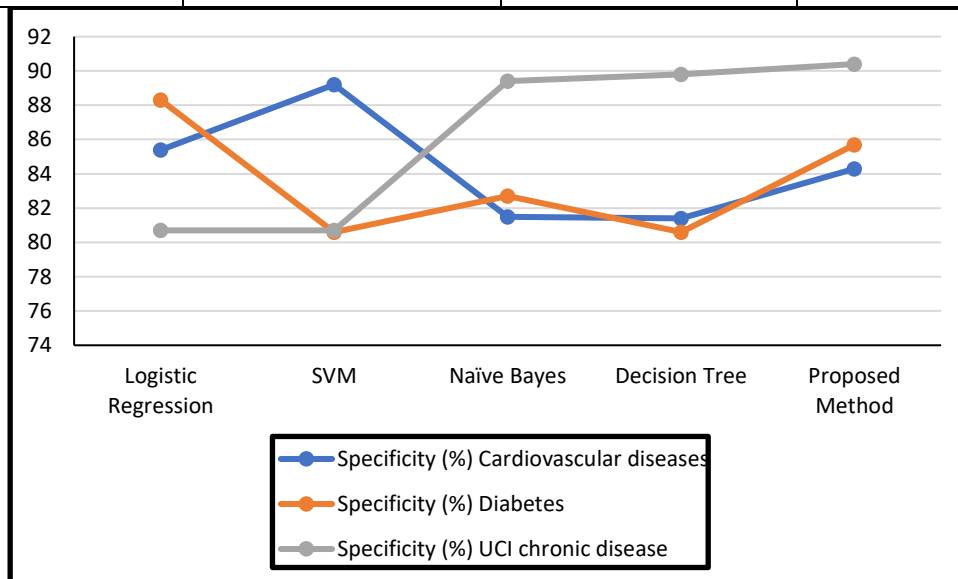


Fig 3: Specificity Performance Comparison of Different classifiers with the projected method.

Figure 1, 2 and 3 are a graphical representations of the experiment's results. The unique ensembling technique

clearly improved the prediction systems' performance. The heart disease prediction method was found to have

an accuracy, sensitivity, and specificity of 87.8%, 88.5%, and 89.3%, respectively.

It was also shown that diabetes prediction systems performed exceptionally well. The results showed a sensitivity of 90.7%, specificity of 85.7%, and accuracy of 92.4%. Prediction systems for chronic renal disease were shown to have an accuracy of 86.4% and a sensitivity of 89.4%. It's clear from these findings that the proposed methodology can be used to create reliable prediction systems for long-term conditions.

5. Conclusion:

Affordable, readily available, and high-quality medical care is still out of reach for the vast majority of Indians. When medical diagnostic and therapeutic aid is delayed in the early stage of disease, morbidity and mortality rise because the disease has progressed further. Over the past two decades, India has seen a startling rise in the death rate from non-communicable diseases. Prediction algorithms for chronic diseases based on machine learning were created in this study. The authors developed an innovative ensembling method. The experimental findings show that the proposed method successfully improves the classifiers' overall efficiency. This study's disease prediction system demonstrated impressive performance and can be effectively used for early diagnosis of chronic diseases. These prototypes will be a technical boon in the fight against the rising death toll from chronic illness. A significant step towards better healthcare would be the implementation of these chronic disease prediction systems in primary care settings. This research will be expanded to include the development of prediction models for additional diseases.

References

- [1] G. Battineni, G. G. Sagaro, N. Chinatalapudi, and F. Amenta, "Applications of machine learning predictive models in the chronic disease diagnosis," *Journal of Personalized Medicine*, vol. 10, no. 2, p. 21, 2020.
- [2] B. Manjulatha and P. Suresh, "An ensemble model for predicting chronic diseases using machine learning algorithms," in *Smart Computing Techniques and Applications*, pp. 337–345, Springer, New York, NY, USA, 2021.
- [3] C.-H. Jen, C.-C. Wang, B. C. Jiang, Y.-H. Chu, and M.-S. Chen, "Application of classification techniques on development an early-warning system for chronic illnesses," *Expert Systems with Applications*, vol. 39, no. 10, pp. 8852–8858, 2012.
- [4] D. Gupta, S. Khare, and A. Aggarwal, "A method to predict diagnostic codes for chronic diseases using machine learning techniques," in *Proceedings of the 2016 International Conference on Computing, Communication and Automation (ICCCA)*, pp. 281–287, IEEE, Greater Noida, India, April 2016.
- [5] M. Chen, H. Yixue, K. Hwang, L. Wang, and L. Wang, "Disease prediction by machine learning over big data from healthcare communities," *Ieee Access*, vol. 5, pp. 8869–8879, 2017.
- [6] R. Ge, R. Zhang, and P. Wang, "Prediction of chronic diseases with multi-label neural network," *Ieee Access*, vol. 8, pp. 138210–138216, 2020.
- [7] H. MacLeod, S. Yang, O. Kim, C. Kay, and S. Natarajan, "Identifying rare diseases from behavioural data: a machine learning approach," in *Proceedings of the 2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pp. 130–139, IEEE, Washington, DC, USA, June 2016.
- [8] M. A. Myszczyńska, P. N. Ojamies, A. M. B. Lacoste et al., "Applications of machine learning to diagnosis and treatment of neurodegenerative diseases," *Nature Reviews Neurology*, vol. 16, no. 8, pp. 440–456, 2020.
- [9] Preethi and K. Dharmarajan, "Diagnosis of chronic disease in a predictive model using machine learning algorithm," in *Proceedings of the 2020 International Conference on smart technologies in Computing, Electrical and Electronics (ICSTCEE)*, pp. 191–96, IEEE, Bengaluru, India, October 2020.
- [10] J. Wiens and E. S. Shenoy, "Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology," *Clinical Infectious Diseases*, vol. 66, no. 1, pp. 149–153, 2018.
- [11] S. Swaminathan, K. Qirko, T. Smith et al., "A machine learning approach to triaging patients with chronic obstructive pulmonary disease," *PLoS One*, vol. 12, no. 11, ArticleID e0188532, 2017.
- [12] Z. Wang, J. W. Chung, X. Jiang, Y. Cui, M. Wang, and A. Zheng, "Machine learning-based prediction system for chronic kidney disease using associative classification technique," *International Journal of Engineering & Technology*, vol. 7, pp. 1161–1167, 2018.
- [13] G. Roth et al., "Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015", *Journal of the American College of*

- Cardiology, vol. 70, no. 1, pp. 1- 25, 2017. Available: 10.1016/j.jacc.2017.04.052
- [14] P. Dhillon et al., "The burden of cancers and their variations across the states of India: the Global Burden of Disease Study 1990–2016", *The Lancet Oncology*, vol. 19, no. 10, pp. 1289- 1306, 2018. Available: 10.1016/s1470-2045(18)30447-9 [Accessed 28 May 2021].
- [15] C. K and B. Thomas, "Cancer trends and burden in India", *The Lancet Oncology*, vol. 19, no. 12, p. e663, 2018. Available: 10.1016/s1470-2045(18)30839-8 [Accessed 28 May 2021].
- [16] S. Varughese and G. Abraham, "Chronic Kidney Disease in India", *Clinical Journal of the American Society of Nephrology*, vol. 13, no. 5, pp. 802-804, 2018. Available: 10.2215/cjn.09180817
- [17] Kumar and A. Pathak, "A machine learning model for early prediction of multiple diseases to cure lives," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 6, pp. 4013–4023, 2021.
- [18] Kalaiselvi, "Diagnosing of heart diseases using average K-nearest neighbor algorithm of data mining," in *Proceedings of the 2016 3rd International Conference on Computing for sustainable Global Development (INDIACom)*, pp. 3099–3103, IEEE, New Delhi, India, March 2016.
- [19] Jain and V. Singh, "Feature selection and classification systems for chronic disease prediction: a review," *Egyptian Informatics Journal*, vol. 19, no. 3, pp. 179–189, 2018.
- [20] Ibrahim and A. Abdulazeez, "The role of machine learning algorithms for diagnosing diseases," *Journal of Applied Science and Technology Trends*, vol. 2, no. 1, pp. 10–19, 2021.
- [21] V. D. Soni, "Chronic disease detection model using machine learning techniques," *International Journal of Scientific & Technology Research*, vol. 9, no. 9, pp. 262–266, 2020.
- [22] P. Ghosh, S. Azam, A. Karim, M. Jonkman, and M. D. Z. Hasan, "Use of efficient machine learning techniques in the identification of patients with heart diseases," in *Proceedings of 2021 the 5th International Conference on Information System and Data Mining*, pp. 14–20, Silicon Valley USA, May 2021.
- [23] Y. Chang and X. Chen, "Estimation of Chronic Illness Severity Based on Machine Learning Methods," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1999284, 13 pages, 2021.
- [24] J. Mishra and S. Tarar, "Chronic disease prediction using deep learning," in *Proceedings of the International Conference on Advances in Computing and Data Sciences*, pp. 201–11, Springer, Valletta, Malta, April 2020.
- [25] E. Jeong, S. Osmundson, C. Gao, D. R. V. Edwards, M. Bradley, and Y. Chen, "Learning the impact of acute and chronic diseases on forecasting neonatal encephalopathy," *Computer Methods and Programs in Biomedicine*, vol. 211, Article ID 106397, 2021.
- [26] Steven Martin, Kevin Hall, Ana Rodriguez, Ana Flores, Ana Silva. *Machine Learning for Emotion Recognition in Educational Settings*. Kuwait Journal of Machine Learning, 2(2). Retrieved from <http://kuwaitjournals.com/index.php/kjml/article/view/186>
- [27] Jahan, K. ., Kalyani, P. ., Sai, V. S. ., Prasad, G. ., Inthiyaz, S. ., & Ahammad, S. H. . (2023). Design and Analysis of High Speed Multiply and Accumulation Unit for Digital Signal Processing Applications. *International Journal on Recent and Innovation Trends in Computing and Communication*, 11(1), 95–102. <https://doi.org/10.17762/ijritcc.v11i1.6055>