International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195

e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

Enhancing and Implementation of ML in Healthcare Sector

Dr. Mahesh Panjwani[1], Sahil Lokhande[2], Rhutwik Gaidhani[3], Sanket Gulhane[4], Sanket Bhute[5], Prathmesh Mathankar[6]

[1]Assistant Professor, Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India.

[2,3,4,5,6]UG Student, Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India

Email Id: *maheshpanjwani@gmail.com[1], sahillokhande987@gmail.com[2], gaidhanirhutwik@gmail.com[3], sanketgulhane13@gmail.com[4], sanketbhute77@gmail.com[5], prathmesh01mathankar@gmail.com[6]*

Abstract

This research paper explores the transformative impact of machine learning (ML) on the healthcare industry, particularly through the emergence of ML-driven health and medical platforms. With the ongoing shift from offline to online shopping, the future of healthcare is envisioned to heavily rely on specialized online doctor consultations and pharmacies. Leveraging ML technologies, these platforms offer personalized services such as real-time video consultations, appointment scheduling, prescription management, and health record maintenance, ultimately prioritizing safety and convenience. This paper elucidates the potential of ML in optimizing patient care, expediting processes, and enhancing decision-making in healthcare. Additionally, it delves into the global trend of online medicine purchasing, highlighting its advantages and examining its impact on the Indian healthcare market.

Keywords: *Appointment Scheduling; E-Pharmacy; Health Prediction; Machine Learning (ML); Online Doctor Consultation.*

1. Introduction

In recent years, the healthcare industry has experienced a profound transformation driven by advancements in technology, particularly in the realm of machine learning (ML). This research paper delves into the remarkable impact of ML on healthcare, focusing on the emergence and implementation of ML-driven health and medical platforms. As society undergoes a significant shift from traditional offline methods to online transactions, the future landscape of healthcare is poised to rely extensively on specialized online doctor consultations and pharmacies. Through personalized services such as real-time video consultations, appointment scheduling, prescription management, and health record maintenance, these platforms prioritize safety and convenience for patients. This paper aims to elucidate the immense potential of ML in optimizing patient care, streamlining processes, and augmenting decision-making within the healthcare sector. Over the years, machine learning (ML) has revolutionized various industries, and its impact on the healthcare sector has been particularly transformative. The rise of ML-driven health and medical platforms marks a significant turning point in healthcare delivery, offering unprecedented levels of personalization and accessibility to patients worldwide. [3&20] As offline transactions give way to online interactions, specialized online doctor consultations and pharmacies emerge as pivotal components of future healthcare systems, all made possible through the leveraging of ML technologies. This paper endeavours to explore the multifaceted roles of ML in enhancing patient care, expediting processes, and improving decision-making within the healthcare landscape, thereby laying the groundwork for a safer, more efficient, and patient-centric healthcare ecosystem. The integration of machine learning (ML) into the healthcare sector represents a paradigm shift in how healthcare services are

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

delivered and accessed. Against the backdrop of a global trend towards online transactions, specialized online doctor consultations and pharmacies emerge as indispensable components of the future healthcare landscape. These platforms offer a wide array of services ranging from real-time video consultations to management, all aimed at prioritizing safety and convenience for patients. This paper aims to delve into the potential of ML in revolutionizing patient care, process efficiency, and decision-making within healthcare, while also examining the implications of the global trend of online medicine purchasing on the Indian healthcare market.

2. Literature Review

- A Machine Learning Approach to Predict Skin Diseases and Treatment Recommendation System, ICSSIT (2023). Author- [1] Suganiya Murugan, S.R. Srividhya, S. Pradeep Kumar, & B. Rubini Kavitha and Murthy's 2019 study uses machine learning techniques, including Naive Bayes, to identify disease types from clinical data, achieving 98.4% accuracy and a GUI for user convenience.
- Derm-NN: Skin Diseases Detection Using Convolutional Neural Network, ICICCS, (2020) Author- [2] Rimi, T. A., Sultana, N., & Ahmed Foysal, M. F. Rimi, Sultana, & Foysal (2020) developed a prototype using CNN for skin condition identification, combining machine learning and photo-processing methodologies to recognize common skin conditions like ulcers, lichen simplex, and dermatitis hand.
- Doctors' Preferences in the Selection of Patients in Online Medical Consultations: An Empirical Study with Doctor–Patient Consultation Data. Healthcare (2022). Author- Rimi, T. A., Sultana, N., & Ahmed Foysal, M. F. The study uses a bipartite graph model and Exponential Random Graph Model to analyze 1404 doctor-patient consultations from a Chinese online medical platform. Results show doctors prioritize moderately ill patients, prefer rural patients, and prefer synchronous communication methods [32].
- Comprehensive Healthcare System Using ML, IJETIR (2021). Author- Abhijit Shendage, Saurabh Gore, Yash Patil, Samyak Vaidya, Aparna Kalaskar. This paper proposes a centralized system using machine learning algorithms to enhance healthcare digitalization. Comparing three algorithms, it found each achieved over 70% accuracy in diagnosing user symptoms. The system aims to unite stakeholders, improve accessibility, and enhance medical assistance [33].
- Virtual online consultations: advantages and limitations (VOCAL) study. Author- [4,5] Greenhalgh T, Vijayaraghavan S, Wherton J, et al. This study examines the use of remote video consultations in healthcare, specifically diabetes and cancer care, within the UK's National Health Service. It uses qualitative methods to understand pros, cons, and ethical aspects, providing insights for academics, providers, policymakers, and patients.
- Face-to-Face with the Doctor Online: Phenomenological Analysis of Patient Experience of Teleconsultation (2022). Author- [6] Grīnfelde, M. The writer examines how COVID-19-driven teleconsultation alters clinical encounters and healing journeys. They explore phenomenology's role, highlighting concerns about virtual consultations lacking physical closeness and empathy. The research aims to determine if teleconsultation maintains these essential aspects. By blending philosophy and qualitative inquiry, it argues for teleconsultation's empathetic potential.
- Advancement of Online Medical Consultation for Future, International Journal of Engineering and Management Research (2022). Author- Abilash.L[7], Danushan.R, Agash.V, Abinaya.T, D.I.De Silva and Dulanji Cooray. This paper delves into the rise of Online Medical Consultation (OMC) via web platforms, highlighting its benefits like convenience, affordability, and improved healthcare access, especially in rural areas. It

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

emphasizes users' positive feedback on its simplicity and effectiveness, showcasing OMC's transformative role in healthcare by linking patients with varied medical professionals.

- Hospital Management System using Web Technology, REVA University, Bengaluru, India (2020). Author – [8] Kotapati Saimanoj, Grandhi Poojitha, Khushbu Devendra Dixit, Laxmi Jayannavar the paper titled "Hospital Management System Using Web Technology" introduces a proposal for a web-based system aimed at substituting paper prescriptions within hospitals. This system advocates for electronic medical management, with the goal of improving patient care efficiency, optimizing doctor schedules, and facilitating convenient access to patient data across the hospital.
- Development of Web and Mobile Based Smart Online Healthcare System, IEEE (2021) Author- [9] M. M. Khan, R. Anwar, F. A. Tanve, D. Shakil, M. Banik and S. K. Gupta This paper delves into the advancement of web and mobile apps in Bangladesh, aimed at enhancing healthcare services for patients. They enable registration, record management, doctor location, and video consultations. Features include medication reminders and online payments, facilitating healthcare accessibility, particularly in remote areas, thus bolstering nationwide healthcare and citizen health in Bangladesh.
- Significance of machine learning in healthcare: Features, pillars and applications, International Journal of Intelligent Networks (2022) Author- [10] Mohd Javaid, Abid Haleem, Ravi Pratap Singh, Rajiv Suman, Shanay Rab This study explores Machine Learning's (ML) impact on healthcare. ML improves physicians' tasks' efficiency and precision, addressing workforce shortages in healthcare. It efficiently handles healthcare data, aiding in sample selection, error reduction, and early epidemic detection. ML streamlines healthcare processes, offering customized treatments, enhancing hospital efficiency, and reducing costs. Its future role includes aiding clinical decisions, illness identification, and personalized treatments, benefiting physicians and hospitals alike.
- Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare, IEEE (2020) Author- [11] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan and A. Saboor The article presents a machine learning system for accurate heart disease diagnosis, utilizing Support Vector Machine, Logistic Regression, and Decision Tree algorithms, as well as feature selection methods like Relief and Least Absolute Shrinkage. Introducing FCMIM for enhanced accuracy and reduced processing time, experiments demonstrate its efficacy compared to existing methods. FCMIM-SVM offers a practical healthcare solution for diagnosing heart disease effectively.
- Machine learning applications in healthcare sector: An overview, Materials Today: Proceedings (2022) Author- [12] Virendra Kumar Verma, Savita Verma the main focus of this paper, application of machine learning (ML) in healthcare, emphasizing its role in analyzing medical data, disease prediction, and safeguarding data privacy. It explores diverse ML algorithms, their applications, methods, and the benefits and challenges they present in healthcare. By addressing the effective use of ML, the paper highlights its significance in improving patient outcomes and medical practices.
- Heart disease prediction using machine learning algorithms, ICCRDA (2020) Author- [13] Harshit Jindal, Sarthak Agrawal, Rishabh Khera, Rachna Jain and Preeti Nagrath In 2021, Harshit Jindal et al. focused on predicting heart conditions using patient medical histories, employing machine learning methods like logistic regression and KNN for classification. Their model, more accurate than naive Bayes, identifies individuals at risk of heart disease, enhancing healthcare efficiency and cost

reduction. The. pynb format presents key insights on heart disease prognosis.

- Effective Heart Disease Prediction Using Machine Learning Techniques. Algorithms 2023. Author- [14] Bhatt, C.M.; Patel, P.; Ghetia, T.; Mazzeo, P.L This study investigates the use of machine learning to forecast cardiovascular disease, a critical aspect for precise diagnosis and treatment. Through the utilization of k-modes clustering with Huang initialization and various models such as random forest, decision tree, multilayer perceptron, and XGBoost, the accuracy of classification is enhanced. Employing an authentic dataset, the models demonstrated impressive performance: decision tree at 86.37% (with CV), XGBoost at 86.87% (with CV), random forest at 87.05% (with CV), and multilayer perceptron at 87.28% (with CV), surpassing alternative methods. Notably, the multilayer perceptron with cross-validation achieved the highest accuracy of 87.28%, underscoring its efficacy in predicting cardiovascular diseases and potentially reducing mortality rates [35].
- Implementation of Machine Learning Model to Predict Heart Failure Disease. Int. J. Adv. Comput. Sci. (2019). Author- [15] Alotaibi, F.S Alotalibi's study delved into the efficacy of machine learning (ML) methods in forecasting heart failure disease, utilizing a Cleveland Clinic Foundation dataset. Employing various ML algorithms such as decision tree, logistic regression, random forest, naive Bayes, and support vector machine (SVM), the research utilized 10-fold cross-validation. Results indicated the decision tree algorithm as the most accurate at 93.19%, with SVM close behind at 92.30%, emphasizing ML's potential in heart failure prediction, notably the decision tree algorithm.
- Heart Disease Prediction Using Machine Learning, Advances in Science and Engineering Technology International Conferences (2022). Author- [16] C. Boukhatem, H. Y. Youssef and A. B. Nassif the study focuses on identifying heart conditions using electronic health records and machine learning. It explores methods like Multilayer Perceptron, Support Vector Machine, Random Forest, and Naïve Bayes for predictions. After data preparation, feature selection, and model assessment with metrics like accuracy, precision, recall, and F1-score, the Support Vector Machine achieved the highest accuracy at 91.67%.
- "Early Diabetic Risk Prediction using Machine Learning Classification Techniques.", IJISRT (2021) Author- [17] Adetunji Olusogo Julius; Ayeni Olusola Ayokunle; Fasanya Olawale Ibrahim. Julius and team employed the Weka platform for analyzing a dataset sourced from the UCI repository, comprising 520 samples with 17 attributes each. Their main goal was to predict early-stage diabetes using machine learning classifiers such as k-NN, SVM, FT, and RFCs. Notably, k-NN yielded the top accuracy of 98%, trailed by SVM at 94%, FT at 93%, and RF at 97%.
- “Machine learning based diabetes classification and prediction for healthcare applications,” Journal of Healthcare Engineering, (2021). Author- [18] U. M. Butt, S. Letchmunan, M. Ali, F. H. Hassan, A. Baqir, and H. H. R. Sherazi. Espino and team used transfer learning and data augmentation to tackle challenges of imbalanced, small training datasets. They explored three neural network architectures, diverse transfer learning methods, augmentation techniques, and loss functions like mixup and generative models. Their designed network, tailored for type 1 diabetes with the OhioT1DM dataset, achieved 95% accuracy.

Table 1 presents an overview of various literature reviews on machine learning (ML) in healthcare. It delves into the broad research on ML's role in disease diagnosis and cancer detection, detailing studied diseases, data sources, algorithms used, evaluation methods, conclusions, and study years.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

This comprehensive summary highlights ML's vast impact in medicine. The below research findings table leads to a significant conclusion on machine learning's impact in medicine. It improves our grasp of ML's role in diagnosing, classifying conditions, and guiding clinical decisions. Comparing results reveals insights, opportunities, challenges, and areas for enhancement in ML and medical diagnostics [36].

Table 1 Research Findings Enables a Notable Conclusion About the Influence and Role of Machine Learning in The Medical Field.

Ref.	Disease	DS	Data source	Algo.	ToA	Result	Year
[19]	Colorectal Cancer	Patients with stage IV colorectal adenocarcinoma	Database BioStudies	LR, DT, GB, lightGBM	CL	LR Accuracy: 91%, DT Accuracy: 89%, GB Accuracy: 84.5%, lightGBM Accuracy: 83.64%	2020
[20]	Rare (CTCs)	Optical and raw-cell microscopy images	Microscopy	CNN	CL	Accuracy: 97%	2020
[21]	Lungs and colon Cancers	Lungs and colon cancer histopathological image	LC25000 dataset from Kaggle	CNN	CL	Accuracy: 96.33%	2021
[22]	Patient's diagnosis	MRI and CT	Private medical center "HT Medica"	SVM, RF, CNN, BiLSTM, NLP	CL	Accuracy: 92.2% (CS=CT), Accuracy: 86.9% (DS=MRI)	2021
[23]	Breast Cancer tumor	Breast cancer tumor gene expression data	The cancer Genome Atlas	KNN, NB, DT, SVM	CL	KNN Accuracy: 92.99%, SVM Accuracy: 96.17%, XGBoost Accuracy: 94.96%, LightGBM Accuracy: 99.86%	2022
[24]	Brain Tumor	CCKS Dataset	CHIP2018, CCKS2019,	CNN	CL	Accuracy: >85%	2022

e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

			and CCKS2020			F1 value: 74.68	
[25]	Breast Tumor	Breast ultrasound image	Local hospital	KNN, SVM, RF, XGBoost, LightGBM	CL	KNN Accuracy: 92.99%, SVM Accuracy: 96.17%, RF Accuracy: 95.08%, XGBoost Accuracy: 94.96%, LightGBM Accuracy: 99.86%	2022
[26]	Brain Tumor	MRI Dataset	Kaggle Website	CNN	CL	Accuracy: 92%	2023
[27]	Heart Disease	UCI	UCI	J48 SVM	CL	J48 Accuracy: 84.35%, SVM Accuracy: 85.03%	2018
[28]	Detecting heart disease	Cleveland		SVM	CL	Accuracy: 92.37%	2020
[29]	Diabetic Prediction using Classification Method	UCI	UCI	SVM, Naives Bayes, Voting classidiers	CL	Accuracy: 83.98%	2020
[30]	Heart Disease Prediction	Kaggle	Kaggle	LR, NB, SVM, KNN, DT, RF and ANN	CL	Accuracy: 90%	2022
[31]	skin lesion classification and Melanoma detection	HAM10000 public dataset	HAM10000 public dataset	DCNN, DenseNet201 neural network model		Accuracy: 78%	2020
DS: Dataset, Algo: Algorithm, ToA: Types of Algorithms, CL: Classification, SVM: Support Vector Machine, BiLSTM: Bidirectional Long Short-Term Memory, NPL: Natural Language Processing, DT: Decision Tree, LR: Logistic Regression, GA: Genetic Algorithm, MLP: Multilayer Perception, GB: Gradient Boosting, lightGBM: Light Gradient-Boosting machine, CNN: Convolutional Neural Network, CT: Computed Tomography, MRI: Magnetic Resonance Imaging, KNN: K-nearest neighbour, NB: Naïve Bayes, CTSs: circulating tumor cells, RF: Random Forest, DCNN: Deep Convolutional Neutral Network, CRF: Conditional Random Field							

3. Research Objective

The research aims to assess how machine learning techniques can enhance health prediction accuracy and prognostic capabilities across medical domains. It focuses on developing predictive models using diverse datasets to detect diseases early, personalize treatment plans, and improve patient outcomes. Additionally, the study investigates integrating medical services like e-pharmacy, live doctor consultation, health diagnosis, and health prediction into a single platform for user convenience, addressing healthcare access challenges, especially in regions like India. Special attention is given to analyzing the benefits and drawbacks of using machine learning algorithms, including Support Vector Machine (SVM) and Random Forest, to predict and manage health conditions like diabetes and heart disease. Furthermore, the research explores the necessity of machine learning in healthcare, examines associated features to strengthen healthcare structures, investigates relevant machine learning frameworks, and delineates significant machine learning applications within the healthcare domain [37].

4. Methodology

4.1. Technology Stack

Python & Django: Python, renowned for its simplicity and clear syntax, boasts a vibrant community supporting numerous libraries and frameworks, particularly for machine learning. With libraries like NumPy, SciPy, pandas, and scikit-learn, Python facilitates data manipulation and analysis. Its adaptability enables seamless integration with healthcare technologies such as databases and web frameworks. Django, a high-level web framework, simplifies web development tasks, including user authentication and database management, while prioritizing security for sensitive healthcare data.

Pandas: Pandas stands out as a powerful Python toolkit designed for handling and analyzing data efficiently. It provides user-friendly data structures such as Data Frames and Series, along with a comprehensive set of tools for tasks like data cleaning, exploration, transformation, and visualization. Moreover, it seamlessly integrates with other popular libraries like NumPy and Matplotlib, enhancing its capabilities even further.

Numpy: NumPy serves as a Python tool tailored for numerical computations, designed to handle sizable arrays and matrices efficiently. Its impressive performance makes it particularly suited for scientific computations such as linear algebra and Fourier transformations. Moreover, its smooth compatibility with various other libraries amplifies its usefulness in tasks involving data analysis and visualization.

Matplotlib: Matplotlib.pyplot serves as a powerful Python library used to generate various types of visual representations, including line plots, scatter plots, bar charts, histograms, and pie charts. Additionally, it seamlessly integrates with NumPy, offers support for multiple backends, and empowers users to design bespoke visualizations.

Seaborn: It is simpler to build statistical visuals for complex datasets and changeable relationships with the help of the Python data visualization module Seaborn. When combined with Pandas, it operates flawlessly, managing aggregation and estimate chores, offering color schemes that can be customized, assisting with regression and correlation analysis, and enabling more customisation using matplotlib methods.

Sklearn: The RandomForestClassifier in scikit-learn employs the Random Forest algorithm for classification, amalgamating predictions from multiple trees to enhance performance, especially for complex datasets. The train_test_split function divides data into training and testing subsets, ensuring fairness and controlling test size. Sklearn.svm.SVC is dedicated to Support Vector Classification, optimizing hyperplanes for maximum margin between classes, offering flexibility with kernels and trade-offs. StandardScaler normalizes features for preprocessing, resilient to outliers, aiding feature selection. Sklearn.model_selection.GridSearchCV fine-tunes hyperparameters through grid search cross-validation, exploring values for generalization performance improvement, necessitating class

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

importation and hyperparameter dictionary definition..

HTM: The common markup language for building websites and online apps is HTML. It defines components like headers, paragraphs, photos, links, and forms using a system of tags and attributes, giving a webpage its structure and content. Web browsers parse HTML documents and display the material in accordance with the guidelines included in the HTML code.

CSS: CSS is a language primarily utilized to dictate how HTML documents should appear and be arranged. It empowers web developers to define various styles, including color schemes, typography, spacing, and placement, thereby tailoring the visual aspects of web pages and maintaining uniformity across diverse pages or devices. CSS operates by linking styles with HTML elements via selectors and property-value combinations, influencing how content appears on web browsers. The most recent iteration, CSS3, brought forth enhancements like gradients, transitions, animations, and responsive design strategies, enriching the creation of captivating and dynamic web interfaces.

JavaScript (JS): JavaScript, a high-level interpreted language, enriches web pages with dynamic features and interactivity, such as form validation, DOM manipulation, and event handling. Its execution in web browsers eliminates the need for server-side processing, facilitating user interactions and the development of interactive web applications with state management and reusable components.

Web APIs: Web APIs are web interfaces that enable software applications to communicate and interact, allowing developers to access and manipulate data, services, or functionality from web servers or third-party services.

4.2. Machine Learning

Machine learning (ML) is a branch of artificial intelligence (AI) and computer science that focuses on the using data and methods that allow AI to replicate human learning processes, thereby increasing its accuracy. Machine learning algorithms make predictions or classifications based on input data. They use an error function to evaluate the model's accuracy, comparing it to known examples. If the model fits better, weights are adjusted to reduce discrepancies between the model estimate and known examples, repeating the iterative process until accuracy is achieved. Our diabetes and heart disease prediction system leverages machine learning algorithms to analyze patient data, including key indicators such as blood sugar levels, cholesterol levels, BMI, and lifestyle factors. Through careful model selection, training, and evaluation, our system can accurately predict the likelihood of diabetes and heart disease in individuals. By deploying these models, healthcare providers can identify at-risk patients early, enabling proactive intervention and personalized healthcare management strategies.

4.2.1. Support Vector Machine

SVM, or Support Vector Machine, is a highly favored technique in Supervised Learning for both Classification and Regression tasks. While primarily used for Classification, it's adaptable to Regression challenges. SVM aims to create an optimal decision boundary, or hyperplane, in an n-dimensional space to separate classes effectively, aiding future data categorization. It identifies crucial support vectors, hence earning its name. Refer to Figure 1 for a visual depiction of classifying two categories using a decision boundary.

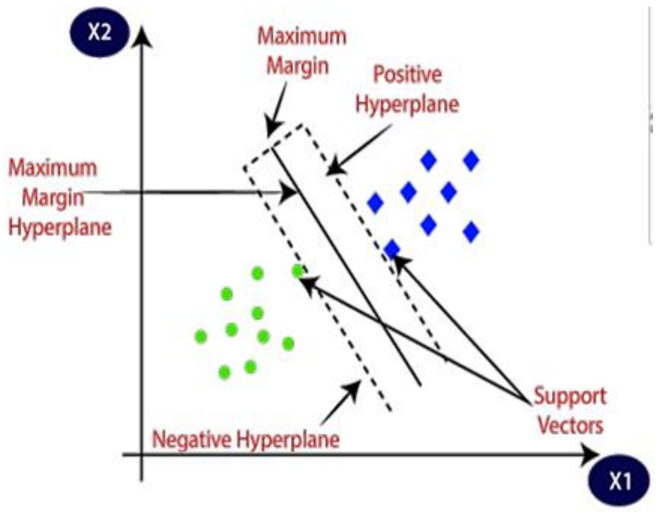

Figure 1 **Analysis of SVM Algorithm**

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

Support Vector Machines (SVMs) excel in discerning intricate data associations, reducing the need for manual adjustments. Ideal for small datasets with numerous features, they surpass other algorithms in handling complex information, yielding more precise results. In medical regression and classification, SVMs aim for maximum class margin and minimal errors, dubbed "Maximum Margin Classifiers." Operating on the Structural Risk Minimization Principle, SVMs establish risk boundaries, utilizing kernel techniques for nonlinear classification. By implicitly mapping inputs into high-dimensional spaces, SVMs ensure clear category separation, optimizing hyperplanes to minimize misclassification while maximizing distance. Given labelled training data as data points of the form M= {(x1, y1), (x2, y2), ..., (xn,yn)} Where yn=1/-1, a constant that denotes the class to which that point xn belonged variable n denotes the number of samples. Each xn denotes a p-dimensional real vector. The SVM classifier operates by converting input vectors into decision values and then performing classification using a suitable threshold. It divides or separates a hyperplane to visualize the training data effectively. This process can be described through the equation: wT.x + b = 0, where 'w' represents the weight of the vector perpendicular to the hyperplane, 'x' denotes the feature vector of a data point, and 'b' stands for the bias term. The objective is to minimize ½||w||2 while ensuring that for all data points, the expression yi(wTxi + b) is greater than or equal to 1, where 'yi' signifies the class label of the data point. The support vectors, or data points that are closest to the hyperplane, are what determine the margin. The margin is maximum in relation to the hyperplane when the support vectors are correctly classified and uniformly spaced. SVMs are a very powerful tool for binary classification since they look for a hyperplane that maximizes the separation between classes. Additionally, by using this method, SVMs become very efficient in a variety of real-world circumstances and resistant to outliers is Figure 2.

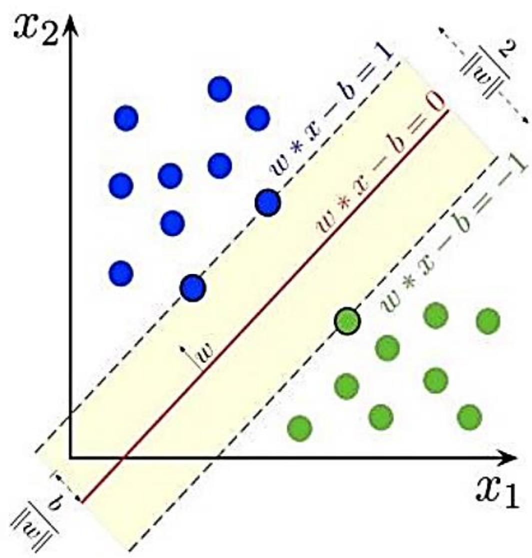

Figure 2 **Maximum Margin Hyperplanes for Svm Trained with Samples from Two Classes**

4.2.2. Random Forest

The Random Forest algorithm is widely used in machine learning for supervised learning tasks. It handles both Classification and Regression problems by employing ensemble learning principles. This involves combining multiple classifiers to tackle complex problems and enhance model performance. Random Forest comprises several decision trees built on different subsets of the dataset. By averaging their predictions, it enhances predictive accuracy. Unlike single decision trees, Random Forest aggregates predictions from each tree and selects the final output based on majority voting. Increasing the number of trees in the forest boosts accuracy and mitigates overfitting issues. Figure 3 illustrates the workings of the Random Forest algorithm. Random Forest employs bootstrapping to build numerous decision trees, creating subsets of training data by random selection with replacement. Each tree is trained on one subset. Additionally, it introduces unpredictability by randomly selecting feature subsets at each node, reducing correlations between trees.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

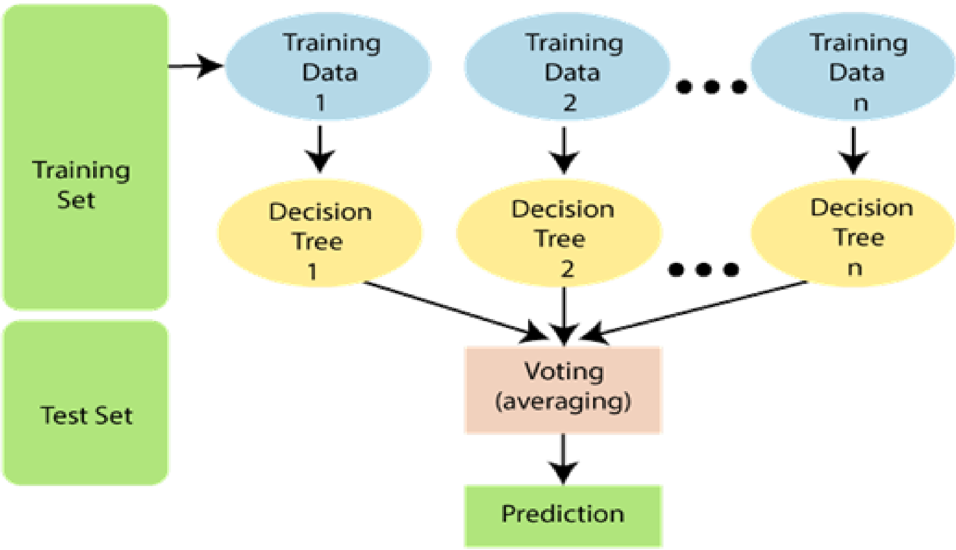

Figure 3 Process of Random Forest Algorithm

Aggregating predictions through majority voting for classification and averaging for regression tasks, it combats overfitting, offering stable forecasts for noisy or high-dimensional data. This approach enhances robustness compared to individual trees, ensuring smoother predictions. Example: Consider a dataset comprising various fruit images provided to the Random Forest classifier. This dataset undergoes subdivision into subsets distributed among individual decision trees. Throughout training, each tree generates predictions, and upon encountering new data, the Random Forest classifier bases its final decision on the majority of these results. Figure 4 illustrates an example of the Random Forest methodology.

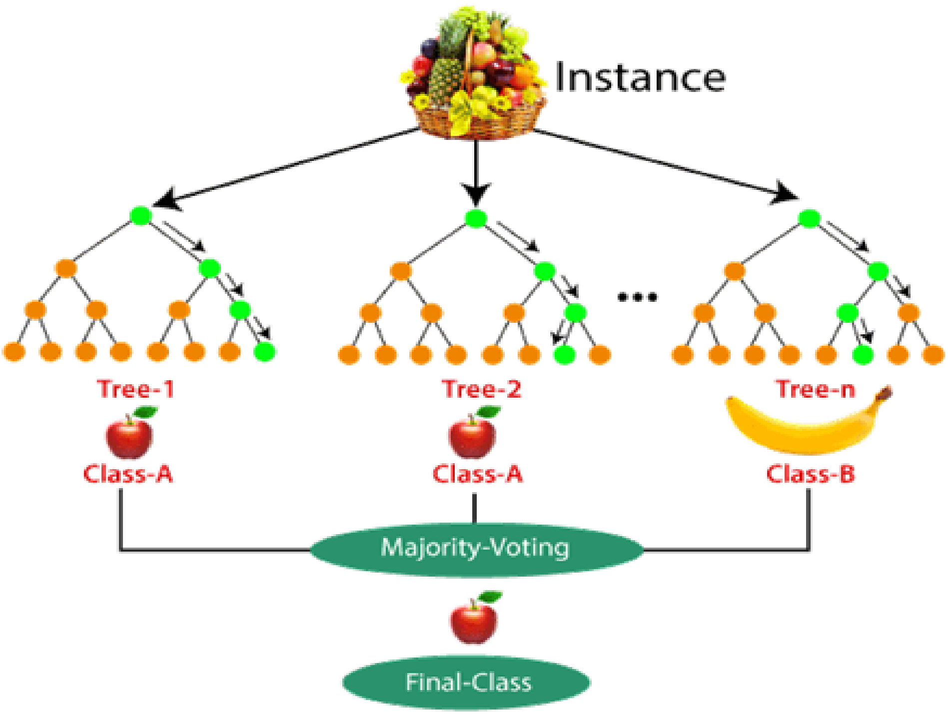

Figure 4 Random Forest Sample Example

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

4.3. Dataset

4.3.1. Heart Prediction System

A meticulously curated dataset of individuals was assembled, considering their cardiac health history alongside other medical factors. Heart disease encompasses various conditions affecting the heart, contributing significantly to mortality rates, especially among middle-aged individuals, as noted by the World Health Organization (WHO). Our dataset comprises medical records from 1329 patients across diverse age brackets, furnishing crucial insights into attributes like age, resting blood pressure, and fasting sugar levels. This information aids in identifying patients diagnosed with heart disease. With 14 medical attributes per patient as shown in Figure 5, our dataset facilitates the classification of those at risk of heart disease and those not at risk. Sourced from the UCI repository, this dataset enables the extraction of patterns indicative of heart disease susceptibility. Split into training and testing subsets, the dataset comprises 1329 rows and 14 columns, each row representing a unique record, with attribute details provided in Table 2. The description of each attribute data in heart disease dataset can be visualized in Figure 6.

Table 2 Various Attributes Used in Heart Disease Prediction Are Listed

Sr. No.	Observation	Description
1.	Age	Age of the user
2.	Sex	Gender of the user
3.	CP	Type of chest pain of user
4.	trestbps	Resting blood pressure value
5.	chol	Cholesterol level of user
6.	fbs	Choosing whether user have fasting blood sugar value or not
7.	restecg	Choosing type of restecg user have
8.	thalch	Maximum heart rate of user
9.	Exang	Choosing whether user have exercise induced angina or not
10.	Oldpeak	Oldpeak value of user
11.	Slope	Choosing type of slope user have
12.	Ca	Number of vessels colored user have
13.	thal	Type of thalassemia user have

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
...	...	...	..	...	...	...	...	...	...	...	...	..	...	...
1323	59	1	1	140	221	0	1	164	1	0.0	2	0	2	1
1324	60	1	0	125	258	0	0	141	1	2.8	1	1	3	0
1325	47	1	0	110	275	0	0	118	1	1.0	1	1	2	0
1326	50	0	0	110	254	0	0	159	0	0.0	2	0	2	1
1327	54	1	0	120	188	0	1	113	0	1.4	1	1	3	0

[1328 rows x 14 columns]

Figure 5 Records in Dataset of Heart Disease

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

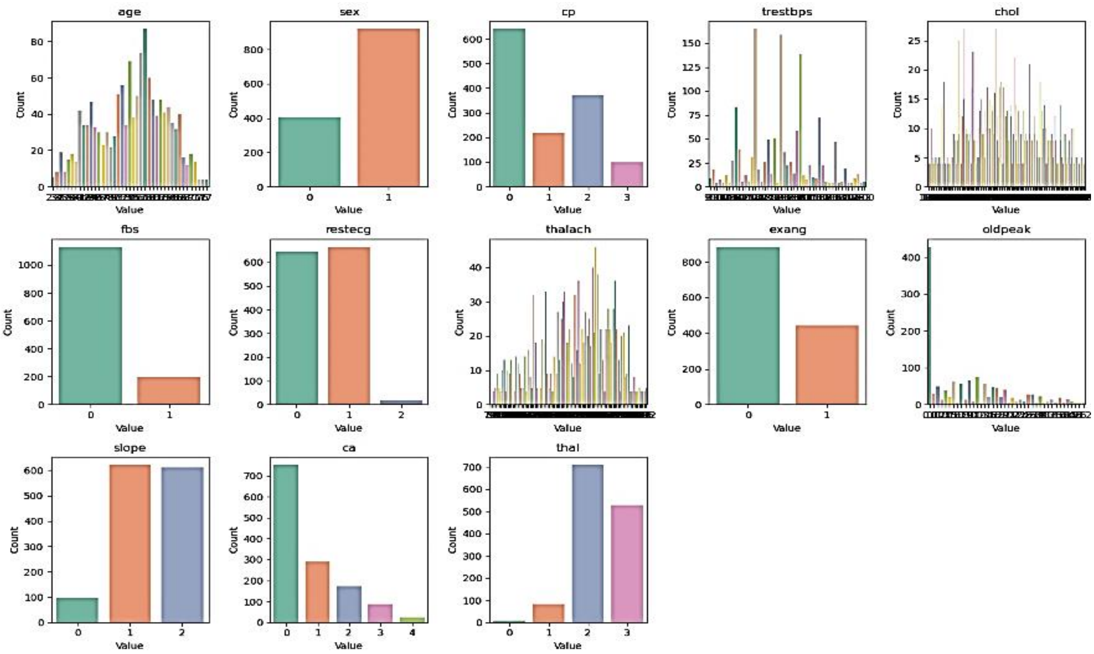

Figure 6 Histogram of Factors Based on Heart Disease Prediction

4.3.2. Diabetic Prediction System

The dataset utilized for the proposed system comprises both numerical and categorical data, allowing predictions for all patients regardless of age or gender. The binary target variable has values "0" indicating a negative diabetes test and "1" indicating a positive result, with more cases in class "0" than "1". A dataset sourced from medical histories of 1537 patients of varying age groups contains 8 medical attributes as shown in Figure 7 and Table 3, aiding in diabetic risk detection and patient classification. Through parameter refinement and manual assessment, variables showcasing optimal discrimination were identified. Attributes in the diabetic disease dataset are detailed in Figures 8 and 9 for visualization. Eight variables—numeric and characteristic—are present: (1) pregnancies, (2) an oral glucose tolerance test's plasma glucose concentration after two hours (HbAlc_level), (3) blood pressure diastolic (mm Hg), (4) Skin thickness, (5)insulin, (6)body mass index(BMI), (7)diabetes pedigree function and (8)age.

Table 3 Various Attributes Used in Diabetic Prediction Are Listed

Sr. No.	Observation	Description
1.	Pregnancies	Pregnancy month of user. Male need to give 0 as input
2.	Glucose	Enter glucose level of user
3.	Blood Pressure	Enter blood pressure of user
4.	Skin Thickness	Take skin thickness of user as input
5.	Insulin	User need to input insulin level
6.	BMI	User gives his/her body mass index
7.	Diabetes Pedigree Function	Enter diabetes pedigree function of user
8.	Age	User need to enter age

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195

e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

	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
...	...	...	...	...	...	...	...	...	...
1531	10	101	76	48	180	32.9	0.171	63	0
1532	2	122	70	27	0	36.8	0.340	27	0
1533	5	121	72	23	112	26.2	0.245	30	0
1534	1	126	60	0	0	30.1	0.349	47	1
1535	1	93	70	31	0	30.4	0.315	23	0

[1536 rows x 9 columns]

Figure 7 Dataset of Diabetes Disease

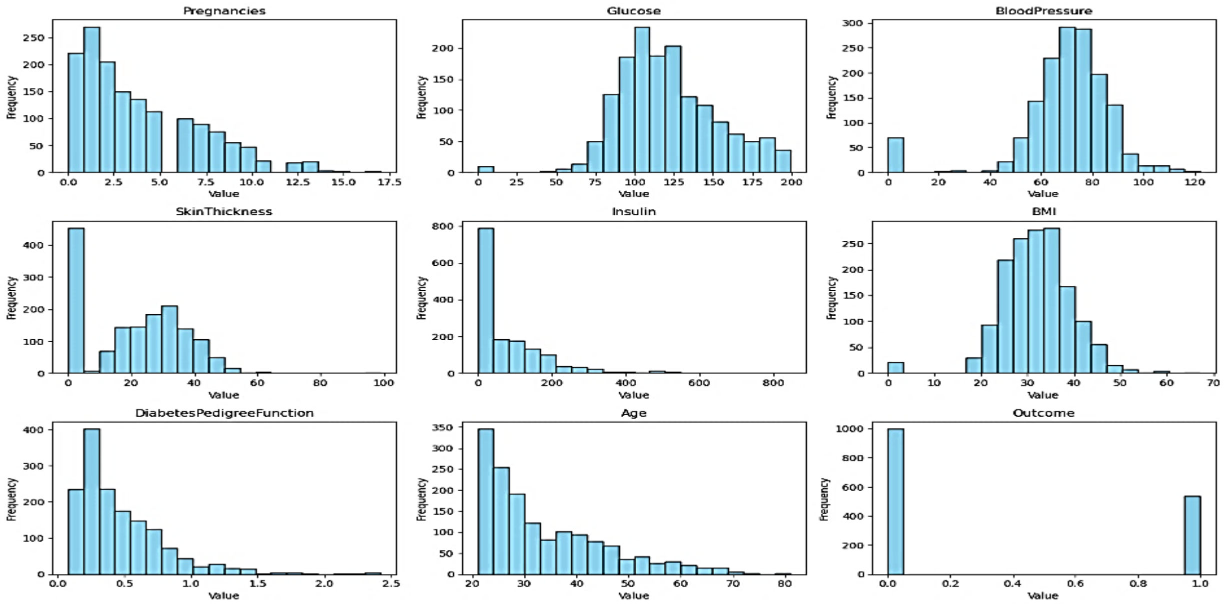

Figure 8 Histogram of Factors Based on Diabetes Prediction

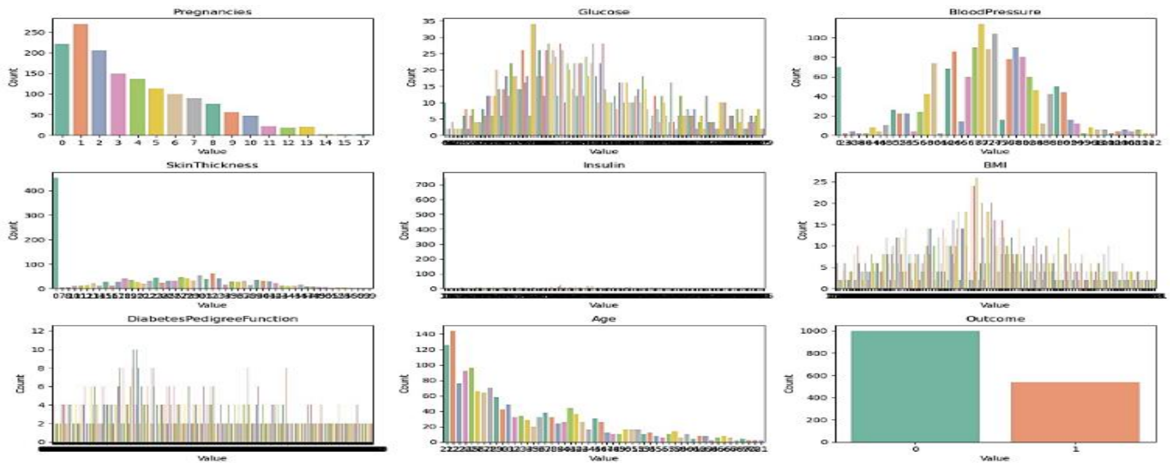

Figure 9 Diabetics Dataset Description (Visualization)

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

4.3.3. Data Processing

Data pre-processing involves readying raw data for machine learning models, a pivotal initial step. Projects often confront unclean or unstructured data. Our model adeptly manages uneven data, enhancing result accuracy. We impute missing values for attributes like Age, Glucose, Smoking, Alcohol, BMI, and Skin Thickness. Subsequently, we normalize the dataset through scaling. Cleaning and formatting data are imperative for operations, thus necessitating data pre-processing.
Real-world data typically has errors, missing numbers, and may be in an unusable format that prevents machine learning models from being applied directly. Preparing the data is a necessary step in order to clean it up and prepare it for a machine learning model, which in turn improves the model's efficiency and accuracy.

4.3.4. Data Manipulation

We obtained our dataset from Kaggle to ensure its relevance and coherence. Our final dataset comprises 1537 rows and 8 columns for diabetes prediction and 1329 rows and 14 columns for heart disease prediction. Initially, we check for missing values to prevent bias, luckily finding none. Then, we rename the data for clarity. Outlier detection involves identifying and removing extreme values, especially in numeric features like age, resting blood pressure, and cholesterol. Before splitting the dataset for training and testing, we encode categorical variables as dummy variables to improve model accuracy.

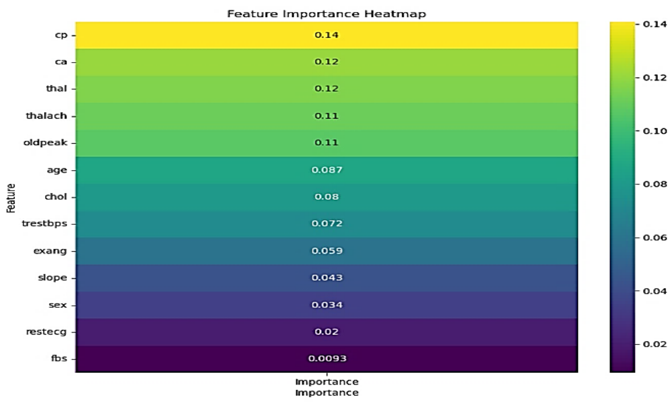

Figure 10 **Heatmap of Heart Disease Along with Attribute**

Finally, we examine correlation patterns between columns and the target value to optimize our dataset for analysis. Figure 10 show the heat map of heart disease and Figure 11 shows heatmap of diabetes disease.

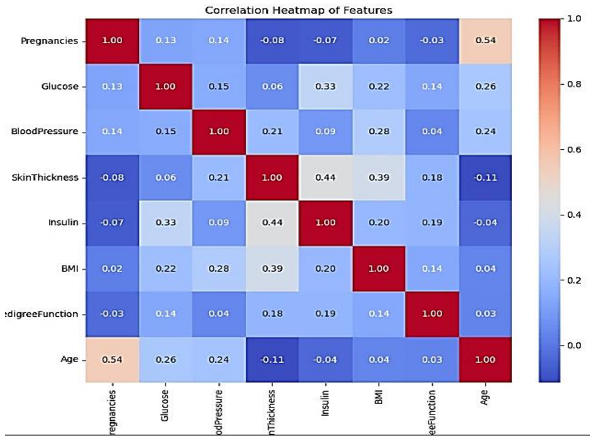

Figure 11 **Heatmap of Diabetes Disease Along with Attribute**

4.3.5. Splitting the Dataset into The Training Set and Test Set

We partition our dataset into a training set and a test set. This study employed an 80% training data and 20% testing data split. This step is critical as it enhances the performance of our machine learning model. Ensuring that both the training and testing datasets are distinct helps the model understand correlations effectively. If the model is well-trained but tested on a different dataset, its performance may decline. Thus, the goal is to create a model that performs consistently well on both training and testing data. The training set comprises data used to train the model, while the test set is used to evaluate the model's predictions.

4.4. Performance Evaluation

We evaluated the model's quality using various performance indicators such as classification accuracy (ACC), precision, recall, F1-score, AUC score, and receiver-operating characteristic (ROC) curve. The performance metrics were determined based on this evaluation process, as illustrated in Figure 12. The confusion matrix, shown in Table 4, helped us assess accuracy, precision, recall, and F1 score during performance evaluation.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

Table 4 Confusion Matrix

	Actual Class	**Positive**	**Negative**
Predicted Class	Positive	True Positive (TP)	False Positive (FP)
	Positive	False Negative (FN)	True Negative (TN)

Various algorithms' performance is assessed using specific metrics. These include True Positive (TP), indicating when a person has a disease and the prediction is positive; True Negative (TN), when a person is disease-free and the prediction is negative; False Positive (FP), when a person is disease-free but the prediction is positive; and False Negative (FN), when a person has a disease but the prediction is negative. Accuracy rate can be determined using TP and TN, while error rates are calculated using FP and FN. True Positive Rate is TP divided by the total number of individuals actually having the disease, and False Positive Rate is FP divided by the total number of individuals without the disease. Precision is determined by TP divided by the total number of individuals predicted to have the disease, and accuracy refers to the total number of correctly classified records.

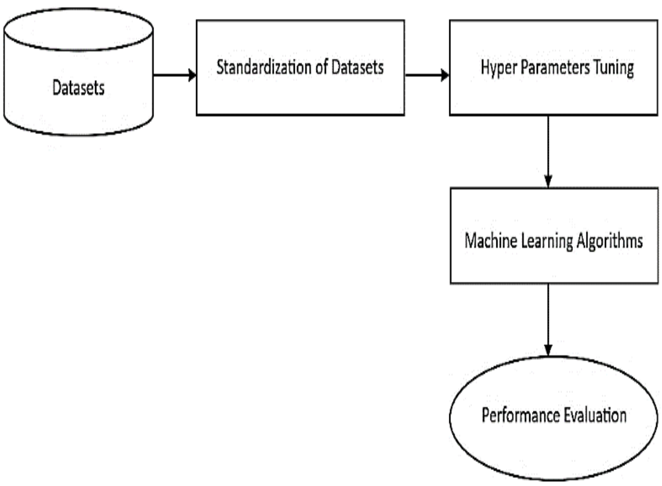

Figure 12 Block Diagram for the Assessment of Performance

Accuracy: Accuracy is determined by the proportion of correct forecasts out of all predictions made. This is mathematically expressed as (1).

$$\text{Accuracy (ACC)} = \frac{true\ positive\ (TP) + true\ Negative\ (TN)}{TP + TN + false\ negative\ (FP) + false\ positive\ (FN)} \quad \ldots. (1$$

Precision: Precision evaluates how many of the positives are identified as meaningful by comparing the actual positives to each positive that the model predicted. In terms of math, it is given by (2)

$$\text{Precision} = \frac{(true\ positives)\ TP}{(true\ positives)\ TP + (false\ positives)\ FP} \quad \ldots. (2)$$

Recall: Recall, also known as Sensitivity, measures the ability of a test to correctly detect positive cases by assessing the proportion of true positives among all the positives within our dataset. This is mathematically represented by equation (3).

$$\text{Recall} = \frac{true\ positives\ \ (TP)}{(true\ positives)\ TP + \ (false\ negatives)\ FN} \quad \ldots. (3)$$

F1-score: The F1 score is a measure that combines recall and precision using a weighted average. It's calculated mathematically as (4).

$$\text{F1 score} = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}} \quad \ldots. (4)$$

The ROC Curve is a graphical representation used to assess the performance of various methods. It shows how sensitivity relates to "1-specificity." This visualization is crucial for understanding algorithm performance, with the area under the curve indicating effectiveness, ranging from 0.5 to 1. Classifier performance is frequently judged using this area.

5. Working

5.1. Implementation of Algorithm

5.1.1. Random Forest

Random Forest, an ensemble method, employs numerous decision trees trained on random data subsets and feature subsets, reducing overfitting and decorrelation. Through majority voting, predictions are made, proving effective for intricate feature relationships and non-linear patterns, such as heart disease risk prediction, and offering insight into key

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

predictive features. In our heart disease prediction system, we've implemented the Random Forest algorithm for its proven effectiveness in tackling classification tasks, particularly within medical diagnostics like forecasting the risk of heart disease. This method, an ensemble learning technique, blends numerous decision trees to yield more resilient and precise predictions. Our project's chosen attributes have been meticulously selected based on their established relevance to forecasting heart disease. These attributes often include factors such as age, gender, cholesterol levels, blood pressure, and various other medical indicators. The Random Forest approach operates by constructing an array of decision trees, each trained on a random subset of the data and utilizing a random subset of features at each stage. This inherent randomness aids in decorrelating the trees and curbing overfitting tendencies. During training, individual decision trees make predictions independently. The final outcome results from aggregating these predictions through majority voting. The Random Forest excels in managing complex attribute relationships and non-linear boundaries, making it perfect for capturing various factors linked to heart disease risk. Integrating Random Forest into our prediction system aims to provide healthcare professionals with precise risk assessments for early detection and timely interventions, thus enhancing patient outcomes and alleviating the burdens of heart disease on individuals and healthcare systems. The heart disease prediction system employing the Random Forest algorithm functions by gathering a dataset brimming with attributes correlated to heart disease risk. Achieving an accuracy rate of 96.59%, this system showcases the transformative potential of machine learning in healthcare by facilitating early detection and intervention for patient welfare. Before model training, the dataset undergoes meticulous preprocessing to optimize performance, following which the model is trained utilizing an ensemble of decision trees. Individual decision trees forecast heart disease risk, collectively forming predictions via majority voting. Model evaluation on the test set includes metrics like accuracy, precision, recall, and F1-score. After training, the model predicts heart disease likelihood using learned patterns, aiding in advancing heart health outcomes in Figure 13.

Pseudo Code of Random Forest:

Precondition: A training set S = (x1, y1), ..., (xn, Yn), features F, and number of trees in forest B.

```
function RANDOM FOREST (S, F)
    H← Ǿ
    for i € 1,, B do
        S(i) ← A bootstrap sample from S
        hi ← RANDOMIZED TREELEARN(S(i), F)
        H← H U{hi}
    end for
    return H
end function
function RANDOMIZED TREELEARN (S, F)
    At each node:
        f← very small subset of F
        Split on best feature in f
    return the learned tree
end function
```

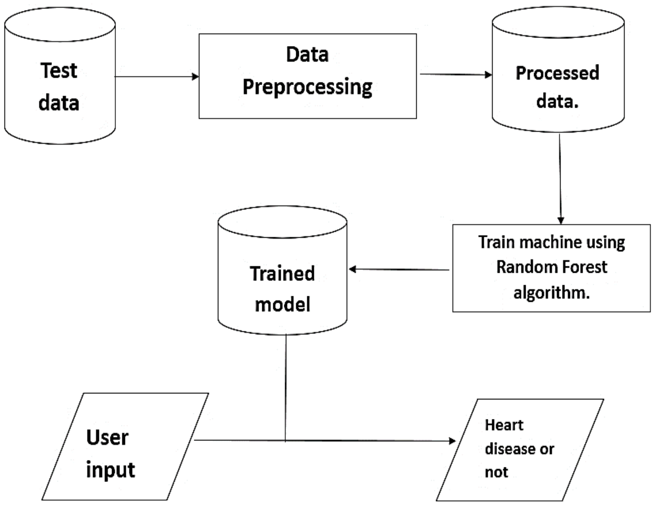

Figure 13 Workflow of Heart Disease Prediction System

5.1.2. Support Vector Machine

In our diabetes prediction system, we harnessed the capabilities of the Support Vector Machine (SVM) algorithm, a potent tool designed for classification tasks, to anticipate the onset of diabetes. SVM proves particularly adept in scenarios featuring two

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

or more classes, functioning by identifying the hyperplane that most effectively segregates data points from distinct classes within a high-dimensional feature space. For our project, we carefully selected attributes known to hold relevance in predicting diabetes, such as glucose levels, body mass index (BMI), age, family history, blood pressure, and insulin levels. Utilizing these attributes as input features, SVM learns to categorize individuals into either diabetic or non-diabetic groups. The core of SVM's strategy is centered on maximizing the margin between classes, aiming to locate the hyperplane that best divides the two groups while minimizing classification errors. SVM's strength lies in its ability to handle non-linearly separable data using kernel functions such as radial basis function (RBF) or polynomial kernels. These kernels transform input features into higher-dimensional spaces, making data linearly separable and enhancing SVM's classification accuracy. SVM's resilience against overfitting is particularly advantageous in medical prediction tasks, ensuring effective generalization to new data. Moreover, its efficiency in managing high-dimensional datasets makes it ideal for scenarios like ours with numerous attributes. Through the implementation of our diabetes prediction system, SVM furnishes valuable insights into the likelihood of diabetes onset for individuals, equipping healthcare professionals with the knowledge needed to make informed decisions regarding patient care and management. By leveraging the capabilities of SVM, our project plays a pivotal role in facilitating early diagnosis and intervention, ultimately enhancing the quality of life for those at risk of diabetes. Utilizing SVM in our diabetes prediction system involves several key steps. We start with meticulous data collection and preprocess the dataset for training suitability. The model then undergoes rigorous training, including hyperparameter tuning to boost performance and reduce overfitting. Its accuracy is later assessed on a separate testing set, ensuring its effectiveness with new data. Achieving an impressive 90.90% accuracy, our system showcases the potential of machine learning in healthcare. This success represents a significant step towards early detection and personalized treatment plans, marking a positive advancement in medical science and patient care is Figure 14.

Pseudo code of SVM:

Input: D= [X, Y]; X (array of input with m features), Y (array of class labels)
Y=array(C) // Class label
Output: Find the performance of the system
function train_svm(X,Y, number_of_runs)
initialize:learning_rate=Math.random();
for learning rate innumber_of_runs
error=0;
for i **in** X
if (Y[i]*(X[i]*w)) <<1 **then**
update: w=w + learning_rate * ((X[i]*Y[i]) *(-2*(1/number_of_runs) *w)
else
update: w=w+learing_rate *(-2*(1/number_of_runs) *w)
end if
end
end

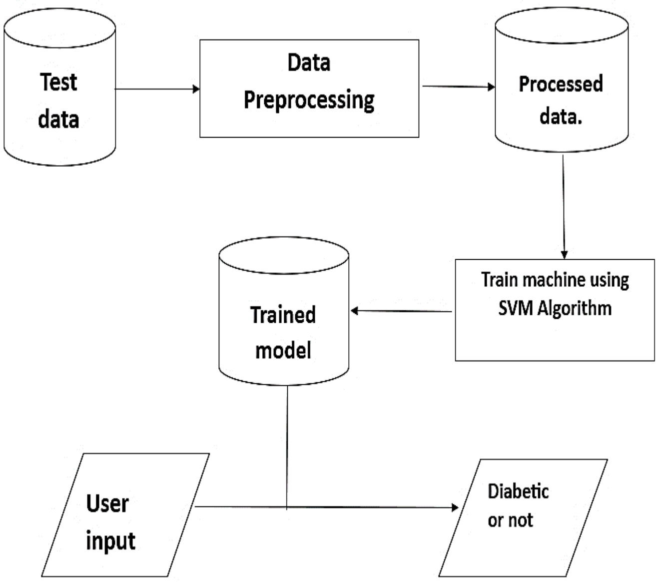

Figure 14 Workflow of Diabetic Prediction System

5.2. System Architecture

This architectural approach harmoniously combines HTML, CSS, and JavaScript in the frontend,

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

supported by Bootstrap for improved design and adaptability. Python, in conjunction with Django on the backend, drives swift development and strengthens resilience. This combination results in smooth user experiences with interactive frontend elements and efficient backend operations. These technologies intertwine to create web applications that are both scalable and easy to maintain, promoting the reuse of code and simplified deployment processes. Through the utilization of HTML, CSS, and JavaScript, alongside Bootstrap, frontend development achieves a consistent design language and responsiveness. Simultaneously, Python and Django empower the backend, ensuring agility in development and robust performance. This holistic strategy guarantees fluid user interactions, dynamic frontend experiences, and streamlined backend functionalities. The integration of these technologies enables the development of web applications that are scalable, maintainable, and user-friendly, with enhanced code reuse and simplified deployment processes. The architectural diagram, as depicted in Figure 15, visually represents this cohesive system.

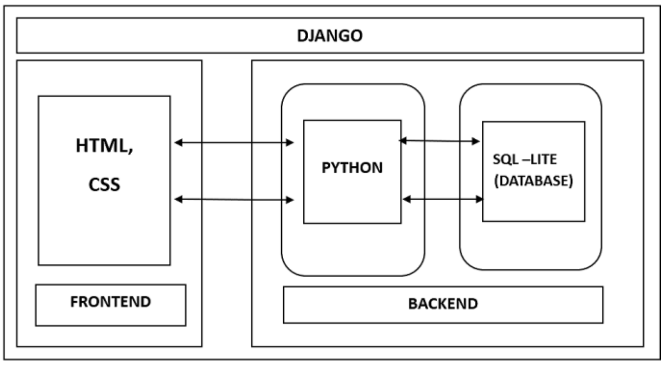

Figure 15 Architecture Diagram

6. Proposed System

The primary objective of this study is to improve healthcare platform and apply machine learning algorithms for the early detection of diseases through a web-based platform. The implementation process utilizes the Waterfall Model, depicted in Figure 16. This model serves as a structured approach to software development, emphasizing sequential stages including requirements analysis, design, implementation, testing, and maintenance. Through the utilization of the Waterfall Model, the project aims to achieve systematic and efficient implementation of the machine learning algorithms within the web application, facilitating accurate and timely disease prediction.

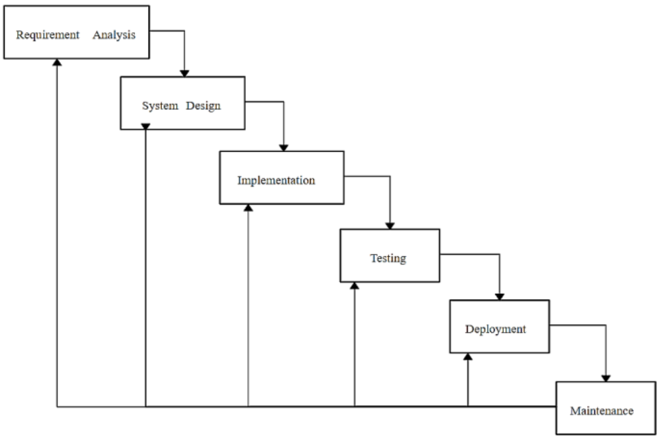

Figure 16 The Waterfall Model

The system proposes a web application platform for the patient and doctors. It provides a summarized view of the medical track record of the patient and gives some insights to the doctor to help him understand the condition of the patient with the help of statistical analysis. The frontend of the system is constructed using HTML, JS, and CSS, which results in a dynamic user interface (UI) that is simple to use and traverse. JavaScript, an object-oriented programming language, is used to offer dynamic interactivity, allowing users to interact with web pages without requiring a page reload. Here, we use python specifically Django framework which comes with integration of SQLite database.

There are three modules in the project are: 1. Patient 2. Doctor 3. Admin

Patients: Patients must register as users. New users can register by providing basic information like name, email, and password. After logging in, they gain access to the dashboard, doctors' profiles, E-pharmacy, and health prediction services. They can then schedule appointments with specific doctors based on their needs. Admins approve or decline appointment requests based on availability.

Doctors: Doctors are able to sign up by providing

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

essential information such as their name, qualifications, specializations, and work history. Following successful registration, doctors can access their accounts by entering their username and password. They can then view patient requests and conduct online consultations if appointments are open.

Admin: The admin serves as a mediator between doctors and patients, overseeing their management and ensuring seamless appointment scheduling. Additionally, the admin verifies and adds new doctors to the database. Within the admin section, all appointment details, including patient and doctor information, are accessible. The website's workflow is shown is Figure 17.

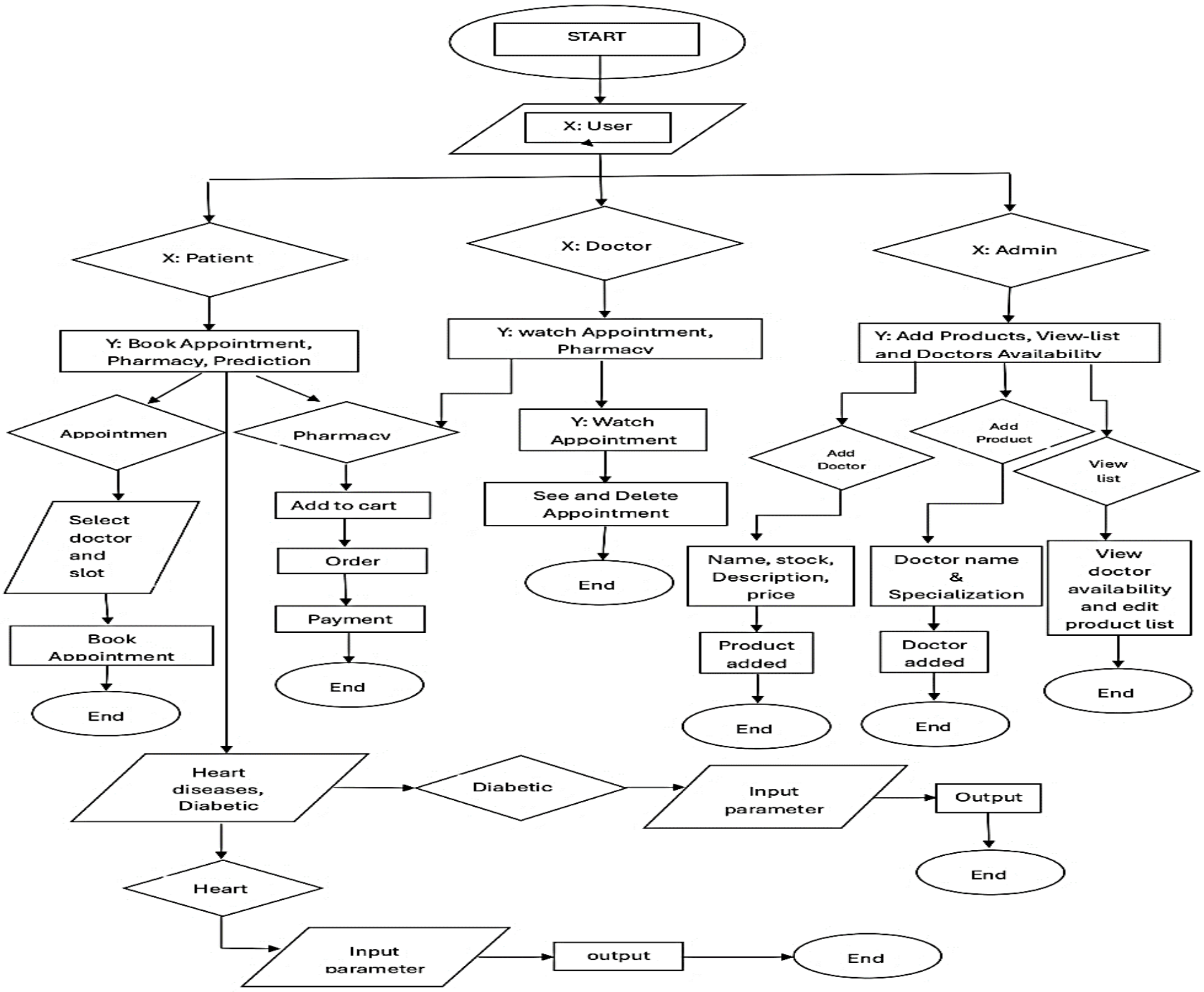

Figure 17 The Workflow of Website

Our interface, crafted with the Bootstrap framework, offers a seamless user experience across various functionalities. Users navigate effortlessly through login, health or disease prediction, online pharmacy access, live doctor support consultation, and appointment booking. The simplicity and intuitiveness of the interface enhance user engagement and satisfaction. Doctors efficiently manage requests and schedules, ensuring timely consultations. Meanwhile, administrators oversee the appointment process, swiftly approving or declining requests as needed. The Bootstrap framework's responsive design ensures accessibility across devices, accommodating users' diverse needs

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

and preferences. With this interface, we prioritize user convenience and functionality, facilitating smooth interactions between patients, doctors, and administrators. Our design ethos centers on simplicity, usability, and efficiency, fostering a productive and supportive healthcare ecosystem.

7. Result and Discussion

The Home page, depicted in Figure 18, acts as users' first stop on the website, providing crucial details and featuring a top navigation bar for interaction options. However, select functions, like services, require registration. Users need to create accounts to unlock these features, encouraging full site exploration.

Figure 18 Home Page

The login page (shown in Figure 19) serves as the gateway for users to access the system by entering their registered username and password. For those without an account, a sign-up button directs them to the registration page. Successful login requires accurate input of both username and password; otherwise, access is denied.

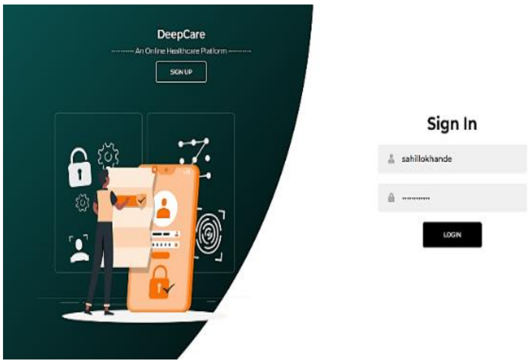

Figure 19 Login

Figure 20 depicts a sign-up form intended for individuals who haven't registered on the website and aim to utilize its services. Accessing system functions necessitates an account, thus absence of one restricts their usage. Users must input their first name, last name, chosen username, email address, and password to complete the sign-up process successfully.

Figure 20 Sign Up

Upon logging into the system, users are presented with three modules: patient, doctor, and admin, as depicted in Figure 21.

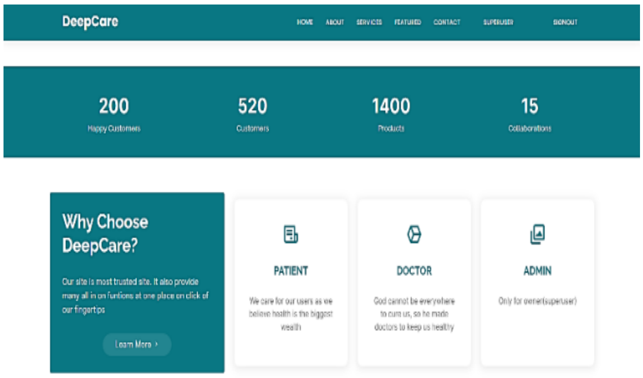

Figure 21 Different Modules

In the Patients module (Figure 22), patients can book consultations with doctors. They enter personal information, choose a time slot, specify the consultation type and ailment category (e.g., dermatology). Next, they pick the appointment date and time (Figure 23) and describe their ailment. The module also offers an online pharmacy (Figure 24) with secure Razor-pay payments (Figure 25).

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195

e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

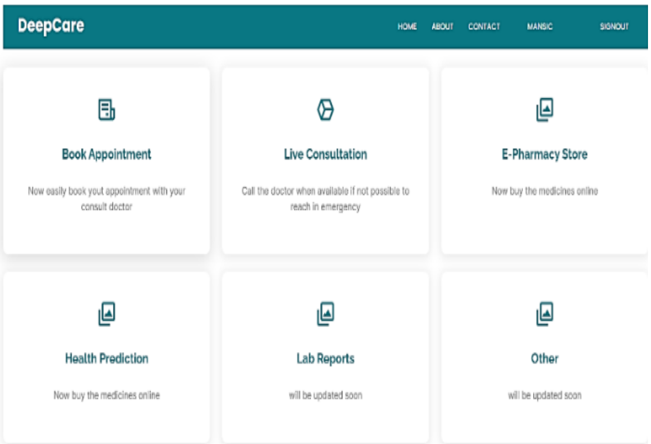

Figure 22 Patient Dashboard

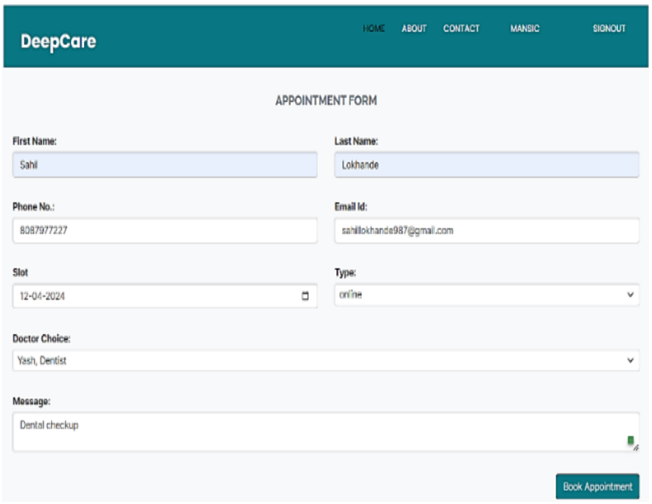

Figure 23 Booking Appointment with Doctor

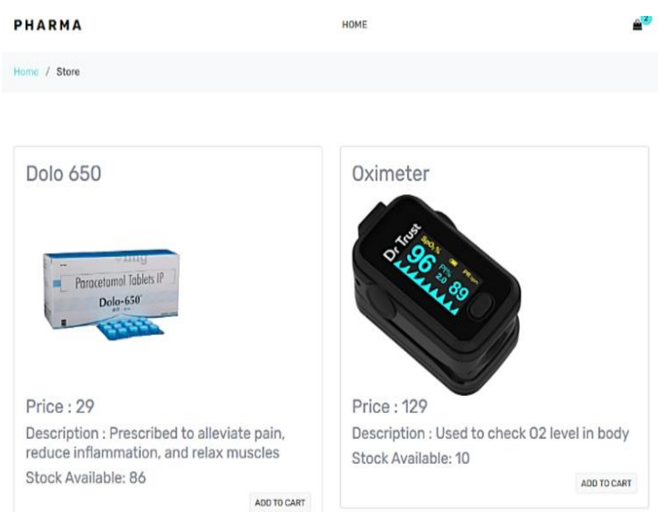

Figure 24 Patient Dashboard

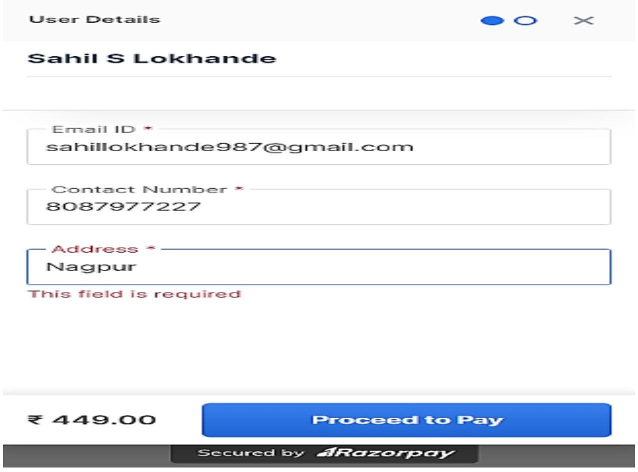

Figure 25 Payment Gateway

Upon accessing the prediction section as shown in Figure 26, two distinct pages emerge: diabetes prediction and heart disease prediction. For diabetes prediction, input parameters (shown in Figure 27) such as glucose, insulin, and blood pressure are required to assess the likelihood of diabetes. Similarly, heart disease prediction necessitates values for parameters like chest pain, age, and cholesterol.

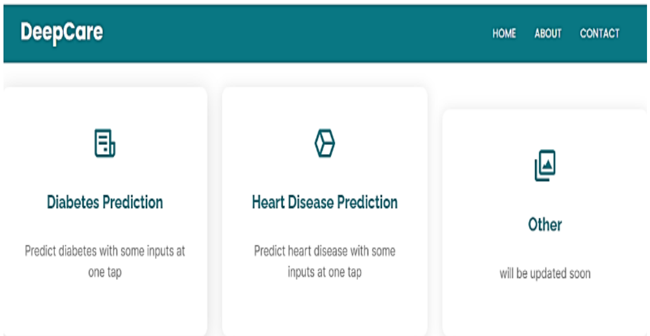

Figure 26 Health Disease Prediction

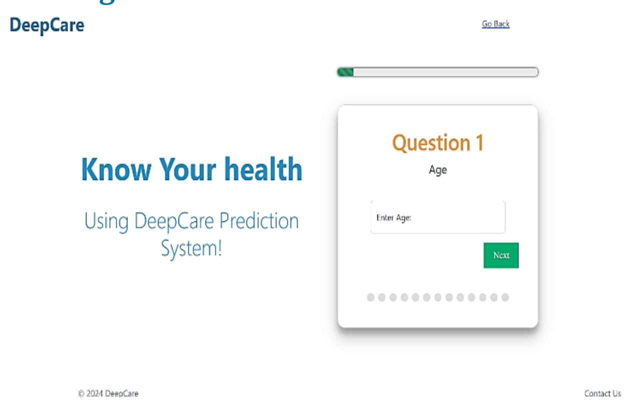

Figure 27 Parameter in Prediction

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

Upon completion of data entry, users are directed to a results page displaying the prediction outcome along with preventive measures. The prediction result page is shown in Figure 28. This user-friendly interface streamlines the process, providing users with valuable insights into their health status and offering proactive steps for prevention.

Figure 28 Prediction Result Page

Users access the "Contact Us" page, depicted in Figure 29, facilitating communication. To reach out, users input their name, email, and message, enabling direct interaction and feedback exchange.

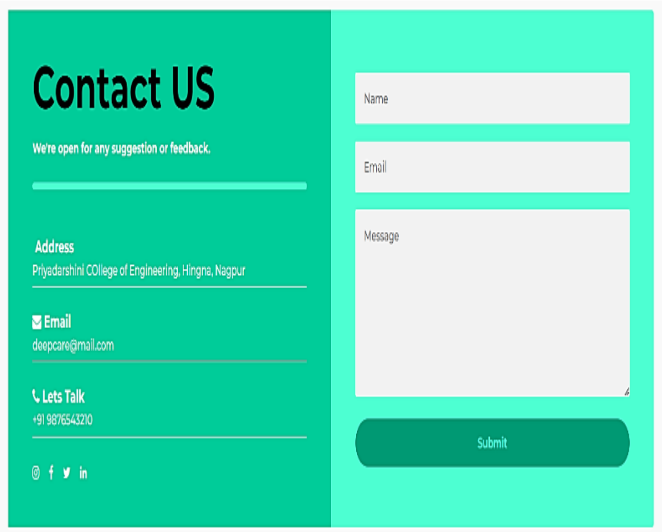

Figure 29 Contact Us Page

Upon doctor login, they access a specialized interface. Figure 30 displays their dashboard with two added pages: one for appointment monitoring (Figure 31) and another for online video consultations (Figure 32). This setup allows doctors to view pending consultations, patient history, manage appointments, and communicate effectively, aiding prompt patient care.

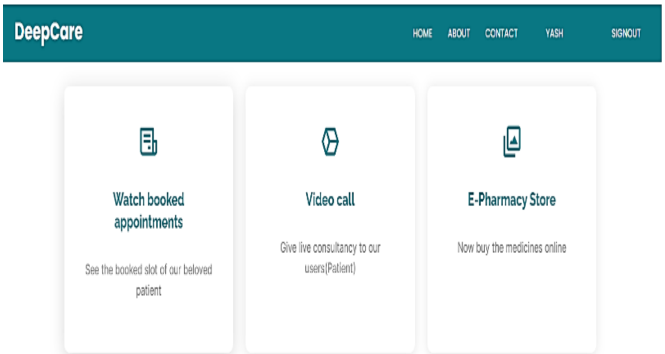

Figure 30 Doctor's Dashboard

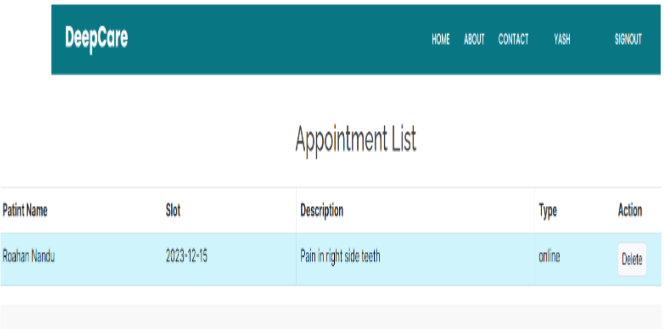

Figure 31 Appointment List

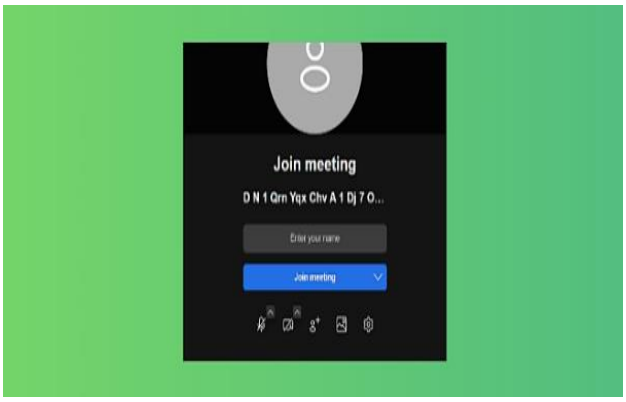

Figure 32 Online Consultation

Upon logging in as an admin, access is granted to a unique interface separate from patient and doctor

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

pages, as depicted in Figure 33. The admin dashboard includes four key sections: adding products, adding doctors, viewing product lists, and viewing order history. Admins have full oversight, allowing them to add new doctors, healthcare products, and medicines, monitor transactions, manage user access, and uphold system integrity and security. Admins play a vital role in regulating access, ensuring system functionality, and maintaining organizational efficiency and user satisfaction.

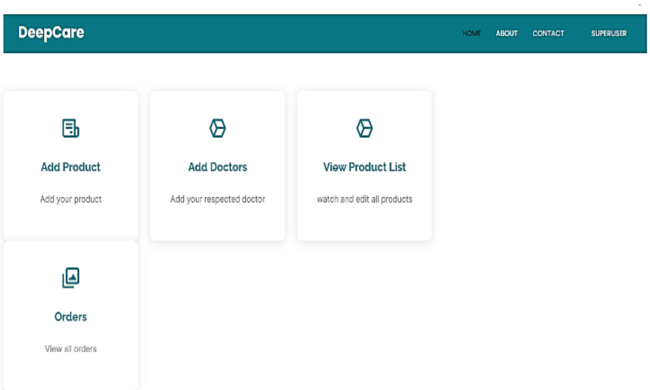

Figure 33 **Admin Dashboard**

Admins can add medicine and healthcare products, inputting essential details like product name, ID, description, stock, price, and a high-quality JPEG or PNG image (Figure 34). This method ensures precise product listings, aiding inventory management and improving the user experience for patients and doctors.

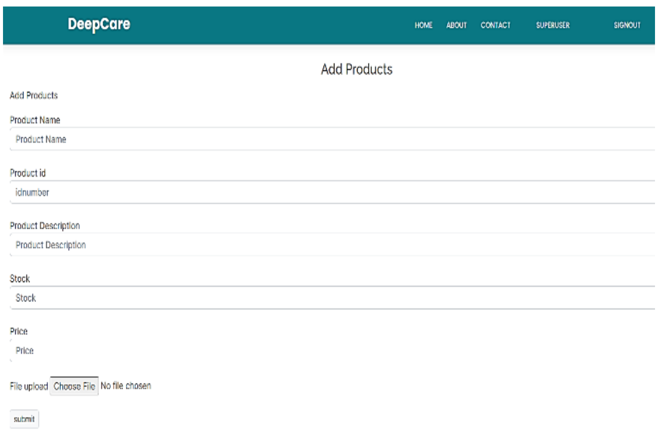

Figure 34 **Add Product**

Admins can add doctors and healthcare products into the system, as shown in Figure 35. When adding a doctor, they must input crucial details like name, specialization, address, fees, and profile picture. After successfully adding the profile, patients can view and consult with the respective doctor.

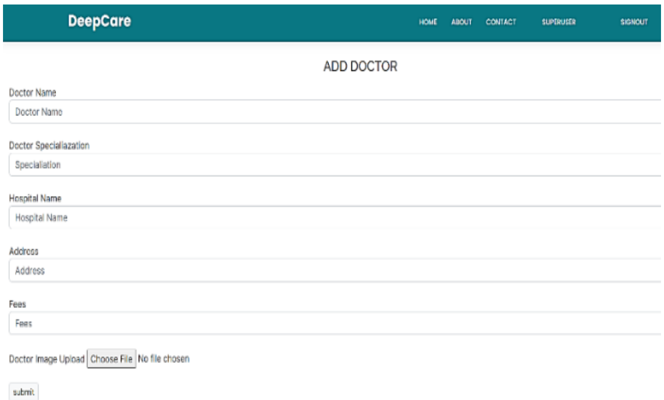

Figure 35 **Adding Doctor's Profile**

Within the admin section, all products added by the admin are listed as shown in Figure 36. Admins can update products by selecting the update button, revealing the complete details of each product for modification or revision.

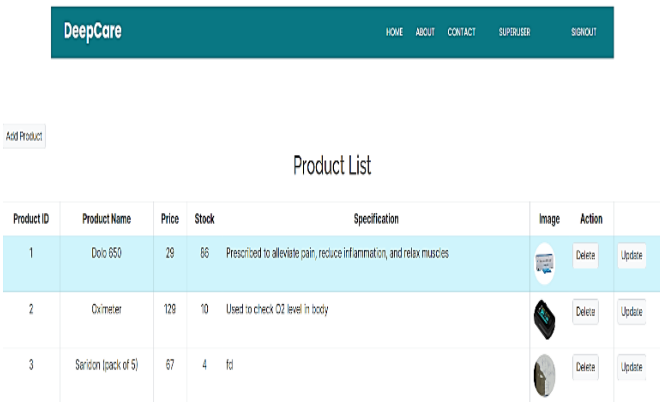

Figure 36 **All Product List**

The screenshots and accompanying explanation effectively detail the proposed system's functionality. Following implementation, rigorous testing methods are applied, adhering to software development standards, to ensure the system's

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

reliability. Finally, after running the algorithms, we calculate the results. There are various parameters used. They are given in table 5.

Table 5 Result of Algorithms

S No.	Model	Accuracy	Sensitivity	F1 Score	Specificity	Precision
1	SVM	0.9090	0.849557	0.872727	0.943589	0.897196
2	RF	0.9659	0.946285	0.972151	0.941647	0.942575

From the chosen algorithms, SVM exhibits an accuracy of 90.90%, sensitivity of 0.84, specificity of 0.94, F1 score of 0.87, and precision around 0.89. Meanwhile, Random Forest based heart disease prediction system achieves an accuracy of 96.59%, sensitivity of 0.94, specificity of 0.94, F1 score of 0.97, and precision approximately 0.94.
Our current project predicts heart disease risk using the Random Forest algorithm and diabetes risk using SVM. We allocate 80% of the dataset for training and 20% for testing. This approach enables accurate model assessment and validation, enhancing the reliability of our predictions for patient health outcomes.

8. Future Enhancement

Upcoming tasks involve integrating a Chatbot and Pathology Center into the system, enhancing it into a comprehensive medical platform for patients. A calendar sync feature will be introduced to send medication reminders to patients. Additionally, the system will accommodate scheduling follow-up appointments. Further enhancements will include home diagnostic tests, personalized health monitoring, and integration with wearable devices for real-time health tracking. The integration of wearable devices aims to provide continuous monitoring and early detection of health issues, such as heart disease.

Conclusion

Through the innovative application discussed in this research, patients can promptly access vital health information without the need for in-person doctor visits, thereby saving both patient and physician valuable time. This approach empowers patients to manage common ailments independently, contributing to a more efficient healthcare system. However, it's imperative to acknowledge the study's limitations, notably its cross-sectional design, which inherently imposes certain constraints. Additionally, the relatively small sample size may limit the generalizability of the findings. Nevertheless, the manuscript introduces robust machine learning techniques, including the Random Forest algorithm for heart disease prediction and the Support Vector Machine (SVM) algorithm for diabetes prediction. These predictive models demonstrate remarkable accuracy in identifying heart disease risk and predicting diabetes onset based on various patient parameters and symptoms. Timely detection of diabetes is crucial for effective treatment, yet many individuals remain unaware of their diabetic status. Therefore, this research underscores the significance of early disease prediction and prevention strategies. Furthermore, it lays the foundation for future endeavours aimed at refining these predictive models to construct more precise and universally applicable tools for early-stage disease risk assessment. Moreover, successful prediction of both diabetes and heart disease enables healthcare providers to proactively allocate resources and tailor treatment plans to individual patient needs, thereby optimizing cost-effectiveness and improving recovery outcomes. By leveraging predictive analytics, healthcare systems can enhance resource management, minimize unnecessary expenses, and streamline overall healthcare delivery efficiency. The integration of predictive models for multiple diseases represents a significant shift in disease management, underscoring the transformative potential of technology in healthcare. With accurate disease prediction, healthcare organizations can prioritize preventive measures, allocate resources effectively, and ultimately enhance patient care and outcomes, ushering in a proactive and personalized healthcare paradigm. In conclusion, this project not only contributes to the burgeoning field of ML implementation in healthcare but also offers a practical and user-friendly solution with tangible benefits for the target community. Acknowledging the study, provides valuable insights for future research directions, encouraging ongoing exploration of innovative approaches to enhance accuracy, inclusivity, and usability in healthcare platforms. By addressing these challenges, we can continue to advance the field and ultimately improve healthcare outcomes for all.

Acknowledgment

It is our pleasure to acknowledge our sincere thanks with a deep sense of gratitude towards our project guide “Dr. Mahesh Panjwani”, Assistant Professor, Computer Technology Department for his continuous support in conducting this dissertation work. He has wholeheartedly helped us in this endeavour at all stages of this work. We are thankful to Dr. S. A. Dhale, Principal, Priyadarshini College of Engineering, Nagpur, for providing the facilities at the institute. We thank Dr. N.M. Thakare, Professor and Head of Computer Technology Department of Priyadarshini College of Engineering, Nagpur. We here with express our immense thanks to “Dr. Mahesh Panjwani”, Assistance Professor, Computer Technology Department, Priyadarshini College of Engineering, Nagpur for giving us suggestions and co-ordination with us from time to time as the project in charge. We also take the opportunity to thank all, who have directly or indirectly extended help and encouragement in executing this project.

Reference

[1]. Suganiya Murugan, S.R. Srividhya, S. Pradeep Kumar, & B. Rubini, (2023), “A Machine Learning Approach to Predict Skin Diseases and Treatment Recommendation System”. 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT).

[2]. Rimi, T. A., Sultana, N., & Ahmed Foysal, M. F. (2020). “Derm-NN: Skin Diseases Detection Using Convolutional Neural Network”. 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)

[3]. Lu, Y.; Wang, Q. Doctors’ Preferences in the Selection of Patients in Online Medical Consultations: An Empirical Study with Doctor–Patient Consultation Data. Healthcare 2022, 10, 1435.

[4]. "Comprehensive Healthcare System Using ML", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.8, Issue 5, page no. g41-g47, May-2021

[5]. Greenhalgh T, Vijayaraghavan S, Wherton J, et al. Virtual online consultations: advantages and limitations (VOCAL) study. BMJ Open 2016;6: e009388.

[6]. Grīnfelde, M. Face-to-Face with the Doctor Online: Phenomenological Analysis of Patient Experience of Teleconsultation. Hum Stud 45, 673–696 (2022).

[7]. Abilash.L1, Danushan.R2, Agash.V3, Abinaya.T4, D.I.De Silva5 and Dulanji Cooray6 : Advancement of Online Medical Consultation for Future, International Journal of Engineering and Management Research e-ISSN: 2250-0758 | p-ISSN: 2394-6962 Volume-12, Issue-5 (October 2022)

[8]. Kotapati Saimanoj, Grandhi Poojitha, Khushbu Devendra Dixit, Laxmi Jayannavar “Hospital Management System using Web Technology” 2020, REVA University, Bengaluru, India.

[9]. M. M. Khan, R. Anwar, F. A. Tanve, D. Shakil, M. Banik and S. K. Gupta, "Development of Web and Mobile Based Smart Online Healthcare System," 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 2021, pp. 0365-0370.

[10]. Mohd Javaid, Abid Haleem, Ravi Pratap Singh, Rajiv Suman, Shanay Rab, Significance of machine learning in

healthcare: Features, pillars and applications, International Journal of Intelligent Networks, Volume 3, 2022, Pages 58-73, ISSN 2666-6030.
[11]. J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan and A. Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," in IEEE Access, vol. 8, pp. 107562-107582, 2020.
[12]. Virendra Kumar Verma, Savita Verma, Machine learning applications in healthcare sector: An overview, Materials Today: Proceedings, Volume 57, Part 5, 2022, Pages 2144-2147, ISSN 2214-7853
[13]. Harshit Jindal, Sarthak Agrawal, Rishabh Khera, Rachna Jain and Preeti Nagrath: Heart disease prediction using machine learning algorithms, IOP Conference Series: Materials Science and Engineering, Volume 1022, 1st International Conference on Computational Research and Data Analytics (ICCRDA 2020) 24th October 2020, Rajpura, India
[14]. Bhatt, C.M.; Patel, P.; Ghetia, T.; Mazzeo, P.L. Effective Heart Disease Prediction Using Machine Learning Techniques. Algorithms 2023, 16, 88.
[15]. Alotaibi, F.S. Implementation of Machine Learning Model to Predict Heart Failure Disease. Int. J. Adv. Comput. Sci. Appl. 2019, 10, 261–268.
[16]. C. Boukhatem, H. Y. Youssef and A. B. Nassif, "Heart Disease Prediction Using Machine Learning," 2022 Advances in Science and Engineering Technology International Conferences (ASET), Dubai, United Arab Emirates, 2022, pp. 1-6.
[17]. Adetunji Olusogo Julius; Ayeni Olusola Ayokunle; Fasanya Olawale Ibrahim. "Early Diabetic Risk Prediction using Machine Learning Classification Techniques." Volume. 6 Issue. 9, September - 2021, International Journal of Innovative Science and Research Technology (IJISRT), ISSN - 2456-2165, PP: - 502-507.
[18]. U. M. Butt, S. Letchmunan, M. Ali, F. H. Hassan, A. Baqir, and H. H. R. Sherazi, "Machine learning based diabetes classification and prediction for healthcare applications," Journal of Healthcare Engineering, vol. 2021, Article ID 9930985, 17 pages, 2021.
[19]. Y. Xu, L. Ju, J. Tong, C-M. Zhou, and J, -J, Yang, "Machine Learning Algorithms for Predicting the Recurrence of Stage IV Colorectal Cancer After Tumor Resection," Sci. Rep., vol. 10, no. 1, p. 2519, Feb. 2020.
[20]. S. Wang, Y. Zhou, X. Qin, S. Nair, X. Huang, and Y. Liu, "Label-freedetection of rare circulating tumor cells by image analysis and machine learning," Sci. Rep., vol. 10, no. 1, p. 12226, Jul. 2020.
[21]. M. Masud, N. Sikder, A.-A. Nahid, A. K. Bairagi, and M. A. AlZain, "A Machine Learning Approach to Diagnosing Lung and Colon Cancer using a Deep Learning-Based Classification Framework," Sensors, vol. 21, no. 3, p. 748, Jan. 2021.
[22]. P. Lopez-Ubeda, M. C. Diaz-Galiano, T. Martin-Noguerol, A. Luna, L. A. Urena-Lopz, and M. T. Martin-Valdivia, "Automatic medical protocol classification using machine learning approaches," Computer Methods and Programs in Biomedicine, vol. 200, p. 105939, Mar. 2021.
[23]. J. Wu and C. Hicks, "Breast Cancer Type Classification Using Machine Learning," J. Pers. Med., vol. 11, no. 2, p. 61, Jan. 2021.
[24]. G. Dhiman et al., "A Novel Machine-Learning-Based Hybrid CNN Model for Tumor Identification in Medical Image Processing," Sustainability, vol. 14, no. 3, p. 1447, Jan 2022
[25]. E. Michael, H. Ma, H. Li, and S. Qi, "An Optimized Framework for Breast Cancer Classification Using Machine Learning," Biomed. Res. Int., vol. 2022, pp. 1-18, Feb. 2022

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0195
e ISSN: 2584-2854
Volume: 02
Issue: 05 May 2024
Page No: 1423-1449

[26]. Muis, S. Sunardi, and A. Yudhana, “Comparison Analysis of Brain Image Classification Based on Thresholding Segmentation with Convolutional Neural Network,” Journal of Applied Engineering and Technological Science (JAETS), vol. 4, no. 2, pp. 664-673, Jun. 2023.
[27]. Chaurasia, V. and S. Pal, “Data Mining Approach to Detect Heart Disease”, International Journal of Advanced Computer Science and Information Technology, vol.2, pp.56-66,2018.
[28]. Li, J.P.; Haq, A.U.; Din, S.U.; Khan, J.; Khan, A.; Saboor, A. Heart disease identification method using machine learning classification in e-healthcare. IEEE Access 2020, 8, 107562–107582.
[29]. Sen, Vimal & Gupta, Krishna. (2020). Diabetic Prediction using Classification Method. International Journal of Recent Technology and Engineering (IJRTE). 9. 264-268. 10.35940/ijrte. F9718.079220.
[30]. Manak, Manaswi, Pankaj Kumar, Garima Gupta. Heart Disease Prediction. International Journal of Innovative Science and Research Technology, Volume 7 – 2022.
[31]. Waweru, A. K., Ahmed, K., Miao, Y., & Kawan, P. (2020), “Deep Learning in Skin Lesion Analysis Towards Cancer Detection”. 2020 24th International Conference Information Visualisation (IV).
[32]. Machine Learning Algorithms-based Prediction of Botnet Attack for IoT devices L. C. Usha Maheswari1, P. Sruthi2, Ch. Hamsini2, G. Keerthika2, M. Sreeja2 1Assistant Professor,2UG Scholar,1,2Department of CSE-Cyber Security 1,2Malla Reddy Engineering College for Women (A), Maisammaguda, Medchal, Telangana.
[33]. Online Medical Consultation: A review of literature and practice Ibrahim Al-Mahdi, Kathleen Gray, Reeva Lederman Department of Computing and Information Systems, Health and Biomedical Informatics Centre the University of Melbourne
[34]. A Method for Improving Prediction of Human Heart Disease Using Machine Learning Algorithms Abdul Saboor,1 Muhammad Usman Muhmmad Faisal Abrar ,1Sikandar Ali ,4andNajeeb Ullah 4 ,2Ali
[35]. An, Q.; Rahman, S.; Zhou, J.; Kang, J.J. A Comprehensive Review on Machine Learning in Healthcare Industry: Classification, Restrictions, Opportunities and Challenges. Sensors 2023, 23, 4178.
[36]. S. Samarpita and R. N. Satpathy, "Applications of Machine Learning in Healthcare: An Overview," 2022 1st IEEE International Conference on Industrial Electronics: Developments & Applications (ICIDeA), Bhubaneswar, India, 2022, pp. 51-56
[37]. Firdous, Shimoo1; Wagai, Gowher A.2, Sharma, Kalpana1. A survey on diabetes risk prediction using machine learning approaches. Journal of Family Medicine and Primary Care 11(11): p 6929-6934, November 2022.