IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

DISEASE PREDICTION USING PAST HISTORIC DATA

[1]Jitendra Kumawat, [2]Lakshay Tanwani, [3]Divy Muradia, [4]Smita Bisht

[1,2,3] Final Year B. Tech Poornima Institute of Engineering and Technology, [4] Assistant Professor Poornima Institute of Engineering and Technology

Abstract: Predicting diseases based on past medical history abstracts involves analyzing patterns, trends, and relevant information within a patient's health records. By leveraging advanced data analytics and machine learning, healthcare professionals can identify potential risk factors, genetic predispositions, and common symptoms associated with specific diseases. This predictive approach enables early intervention and personalized healthcare strategies, aiming to prevent or manage illnesses more effectively. However, it's crucial to note that disease predictions are not infallible, and uncertainties may exist. Ethical considerations, patient privacy, and the need for continuous validation of predictive models are essential aspects in the development and application of such tools. Overall, integrating predictive analytics into medical practices holds promise for enhancing preventive care and optimizing treatment plans based on an individual's unique medical history.

***Keywords* - Predictive healthcare, medical foresight, historical health data, analytics-driven predictions, machine-learned insights, risk profiling, genetic markers, symptom patterns, pre-emptive interventions, individualized health strategies, preventive healthcare**

INTRODUCTION

In contemporary healthcare, the integration of predictive analytics and machine learning has paved the way for groundbreaking advancements in disease predictions based on individuals past medical histories. Analyzing electronic health records and historical health data allows for the identification of subtle patterns, genetic predispositions, and early indicators that can anticipate potential health outcomes. The utilization of sophisticated algorithms enables healthcare professionals to discern risk factors, recognize common symptoms, and personalize predictions according to individual characteristics. This paradigm shift towards predictive healthcare holds the promise of early intervention, personalized treatment strategies, and a focus on preventive care.

By understanding the unique health trajectories of patients, predictive models empower healthcare providers to optimize treatments, implement targeted preventive measures, and enhance overall healthcare efficiency. However, the ethical implications of handling sensitive health data and the ongoing validation of predictive models underscore the importance of responsible and careful integration of these technologies. This introduction sets the stage for exploring how the amalgamation of historical health data and predictive analytics is reshaping healthcare, offering a glimpse into a future where healthcare is not merely reactive but proactive, individualized, and focused on preventing diseases before they manifest.

OBJECTIVE.

A disease prediction project based on past data has numerous utility and application areas that can significantly impact healthcare and public health. Here are some of the key objectives:

1. Early Disease Detection: The primary application is the early detection of diseases. By analyzing historical health data, the system can identify individuals at risk of various diseases well before symptoms manifest. This early detection enables timely intervention and treatment, potentially improving patient outcomes and reducing healthcare costs.

2. Personalized Medicine: Disease prediction models can aid in personalized medicine. They can help healthcare providers tailor treatment plans and medications to individual patients based on their genetic predispositions, past health records, and lifestyle factors.

3. Preventive Healthcare: Disease prediction systems can empower individuals to take a proactive approach to their health. By providing risk assessments and recommendations, patients can make informed decisions about lifestyle changes and preventive measures, such as vaccination and screening.

4. Chronic Disease Management: For chronic diseases like diabetes, hypertension, and heart disease, these systems can assist in ongoing disease management. They can predict exacerbations or complications and help patients and healthcare providers manage these conditions more effectively.

5. Public Health Surveillance: On a broader scale, these systems can be used for public health surveillance. By analyzing population-level historical health data, health agencies can predict disease outbreaks and allocate resources for timely interventions.

6. Resource Allocation: Hospitals and healthcare institutions can use predictive models to optimize resource allocation. For example, predicting patient admissions and disease prevalence can help with staffing and resource planning.

7. Clinical Decision Support: Disease prediction systems can serve as clinical decision support tools for healthcare professionals. They can provide risk assessments, suggest diagnostic tests, and recommend appropriate treatments based on a patient's historical data.

8. Telemedicine and Remote Monitoring: In telemedicine, these systems can remotely monitor patients, providing real-time feedback and alerts to healthcare providers and patients. This is particularly useful for managing chronic diseases and remote patient care.

9. Cancer Screening and Diagnosis: Disease prediction models can assist in cancer screening and diagnosis. For instance, they can analyze medical imaging data to detect early signs of cancer, improving the chances of successful treatment.

10. Drug Development and Clinical Trials: Pharmaceutical companies can use predictive models to identify potential participants for clinical trials based on their risk profiles. This can accelerate drug development and increase the chances of finding effective treatments.

11. Mental Health Prediction: These systems are not limited to physical diseases. They can also predict mental health conditions and allow for early interventions, offering a significant benefit in terms of mental health care.

12. Disease Risk Assessment for Insurance and Wellness Programs: Insurance companies and wellness programs can use these models to assess an individual's disease risk. This information can be used to determine insurance premiums or offer wellness incentives.

13. Epidemiological Research: Disease prediction projects can contribute to epidemiological research by providing insights into disease trends, risk factors, and the impact of interventions over time.

In summary, disease prediction based on past data has wide-ranging applications in healthcare, ranging from individual patient care to population-level public health management. These applications can lead to better health outcomes, reduced healthcare costs, and improved overall well-being for individuals and communities.

RESEARCH METHODOLOGY

The methodology section outlines the plan and method that how the study is conducted. This includes the Universe of the study, a sample of the study, Data and Sources of Data, study's variables, and an analytical framework. The details are as follows;

3.1Population and Sample

Predicting diseases based on past data, researchers typically use historical data from a population to identify patterns, risk factors, and predictive indicators associated with the disease. They may then use this information to develop models or algorithms that can predict the likelihood of individuals within the population developing the disease in the future.

3.2 Data and Sources of Data

Data sources can be obtained from various sources such as healthcare institutions, public health agencies, research databases, and commercial data providers. Integration and analysis of diverse datasets using advanced analytical techniques are essential for developing accurate and reliable disease prediction models based on past data. Additionally, ensuring data privacy, security, and ethical considerations are critical throughout the data collection and analysis process.

3.3 Theoretical framework

The theoretical framework for disease prediction based on past data relies on epidemiological principles and statistical methodologies. By analyzing historical health records, genetic information, lifestyle factors, and environmental influences, researchers identify patterns and risk factors associated with the disease. Utilizing predictive modeling techniques such as machine learning or regression analysis, they develop algorithms to forecast the likelihood of disease occurrence in specific populations. This framework integrates multidisciplinary approaches to inform preventive strategies, early intervention, and personalized healthcare. Ultimately, it aims to improve public health outcomes by identifying high-risk individuals and implementing targeted interventions for disease prevention and management.

3.4 Major Problems Identified

Developing a Disease Prediction System based on past data is a complex and challenging endeavor, and it can face several major problems and obstacles. Some of the key issues in such projects include:

1. Data Quality and Availability: Inaccurate or incomplete historical health data can significantly impact the performance of predictive models. Data quality issues, such as missing values or errors, may require extensive cleaning and preprocessing.

2. Data Privacy and Security: Ensuring the privacy and security of sensitive health data is a critical concern. Adhering to data protection regulations and maintaining confidentiality while using patient records is a complex issue.

3. Data Integration: Combining data from diverse sources, including electronic health records, medical imaging, and genetic information, can be challenging. Data integration issues can lead to inconsistencies and hinder model development.

4. Data Imbalance: Some diseases are rare, leading to imbalanced datasets where there are fewer positive cases. Imbalanced data can result in biased models that struggle to predict the minority class effectively.

5. Feature Selection: Identifying the most relevant features or variables from the vast historical data can be difficult. Poor feature selection can lead to overfitting, increased model complexity, or reduced model performance.

6. Model Generalization: Ensuring that predictive models generalize well to unseen data is a major challenge. Models may perform well on training data but poorly on new patient data, which is a significant concern for real-world deployment.

7. Interpretable Models: Many machine learning models, especially deep learning models, are considered "black boxes" that make it difficult to explain why a particular prediction was made. Interpreting model decisions is important for gaining the trust of healthcare professionals and patients.

8. Ethical Considerations: The use of historical health data for prediction raises ethical questions related to informed consent, data ownership, and the potential for bias and discrimination in healthcare decision-making.

9. Real-Time Monitoring and Adaptation: Disease prediction models should be able to adapt to changing health trends and emerging diseases. Developing systems that can continuously update and remain relevant over time is a challenge.

10. Integration into Healthcare Workflow: Implementing predictive models into the existing healthcare workflow, such as electronic health record systems, can be complex. Healthcare professionals may resist or struggle to integrate new technologies into their practices.

11. Regulatory Compliance: Compliance with healthcare regulations and standards, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States, can be burdensome and impact the development and deployment of predictive systems.

12. Limited Data for Rare Diseases: For rare diseases, there may be limited historical data available for training predictive models. This scarcity of data can make it challenging to build accurate models for rare conditions.

13. Patient Heterogeneity: Patients are diverse in terms of genetics, demographics, and lifestyles. Predictive models need to account for this heterogeneity and provide individualized predictions.

Overcoming these challenges requires a multidisciplinary approach involving data scientists, healthcare professionals, ethicists, and regulatory experts. It also necessitates a commitment to data quality, transparency, and continuous evaluation of model performance in real-world clinical settings. Developing robust disease prediction systems that genuinely benefit patients and healthcare providers is a complex and ongoing process.

3.5 Proposed Work

In our exploration of disease predictions based on past medical history within our traditional healthcare system, we have identified several significant challenges that warrant attention and innovation. Firstly, the manual and often fragmented nature of our traditional medical record-keeping poses a substantial obstacle. The reliance on paper-based records or disparate electronic systems hinders the seamless integration of patient data essential for robust predictive analytics.

Moreover, interoperability issues across healthcare facilities limit the holistic view of an individual's medical history. The lack of standardized formats and protocols impedes the efficient exchange of information, diminishing the effectiveness of predictive models that thrive on comprehensive datasets.

Furthermore, a notable challenge lies in the variability of data quality. Inconsistent recording practices, inaccuracies, and incomplete information compromise the reliability of historical health data, directly impacting the precision of predictive analytics. Addressing these data quality concerns is imperative for ensuring the accuracy and trustworthiness of disease predictions.

Additionally, the slow adoption of advanced technologies, such as machine learning and predictive analytics, within our traditional healthcare system is a significant hurdle. Resistance to change, coupled with limited awareness and understanding of these innovative approaches.

Patient privacy concerns also emerge as a critical issue. The sensitive nature of health data necessitates stringent safeguards to protect individual privacy. Striking a balance between utilizing past medical history for predictions and ensuring robust data security measures is a complex yet crucial aspect that demands careful consideration.

In conclusion, the transition towards leveraging past medical history for disease predictions faces formidable challenges within our traditional healthcare system. Addressing issues related to data integration, interoperability, data quality, technology adoption, and privacy protection is imperative to unlock the transformative potential of predictive analytics, ultimately enhancing efficiency and effect. The proposed work aims to overcome the challenges in disease prediction based on past medical history in our healthcare system and move towards more efficient, proactive, and personalized healthcare practices.

To achieve this, we propose the development and adoption of an Integrated Electronic Health Records (EHR) system that will allow seamless data exchange among healthcare providers and ensure a comprehensive repository of patient information.

We also intend to standardize data formats and protocols and implement data quality improvement measures, such as regular audits and training programs, to enhance the reliability of historical health data.

Moreover, we suggest integrating advanced technologies, such as machine learning and predictive analytics, into routine healthcare practices. To make this possible, training programs will be initiated to familiarize healthcare professionals with these tools.

In addition, we recognize the importance of patient privacy and propose initiatives to empower patients. This involves transparent communication about the use of their health data for predictive purposes and obtaining informed consent.

We also propose collaborative efforts with research institutions to conduct rigorous validation studies, ensuring the accuracy and reliability of predictive models across diverse patient populations.

To establish comprehensive policies and legal frameworks concerning the ethical use of predictive analytics in healthcare, we will undertake stringent measures for data security, privacy protection, and responsible implementation.

Finally, we suggest initiating pilot programs in select healthcare facilities to facilitate a smooth transition. This phased approach allows for iterative implementation, enabling real-time feedback and adjustments based on practical experiences and evolving needs.

By undertaking these proposed initiatives, we aim to overcome existing challenges and usher in a new era of healthcare that optimizes patient care and positions our healthcare system as a leader in leveraging innovative technologies for the benefit of public health.

3.5 Methodology

The methodology employed for disease predictions based on past medical history involves a comprehensive and systematic approach, integrating advanced technologies and established research methodologies. The following outlines the key components of the methodology:

1. Data Collection and Integration: This refers to the process of collecting and combining information from various sources into a unified form of information. The aim is to ensure that information is accurate, consistent and complete so that it can be verified and used effectively. This process will involve the use of various tools and technologies such as data warehousing, ETL (Extract, Transform, Load) processes and data integration.

Standardization: Standardizing data formats and ensuring interoperability to overcome challenges related to fragmented and diverse data sources.

2. Data Preprocessing and Cleaning: Quality Assurance: Implementing measures for data quality improvement, including regular audits and validation checks to address inaccuracies and inconsistencies.

Normalization: Standardizing variables and formats to ensure uniformity in the dataset.

3. Feature Selection and Engineering: Identification of Relevant Variables: Determining key features, such as genetic markers, lifestyle factors, and historical diagnoses, through thorough analysis of medical literature and expert consultation.

Feature engineering: Creating new variables or modifying existing variables to increase the predictive power of a model.

4. Machine Learning Models: Algorithm selection: Select the appropriate machine learning algorithm according to the nature of the prediction task (such as classification, regression).

Training and testing: Split the dataset into training and testing parameters to train the model and evaluate its performance.

Model adjustment: Update the model to get the best view.

5. Validation and Evaluation: Cross-Validation: Employing cross-validation techniques to assess model robustness and generalizability.

Metrics: Using relevant evaluation metrics (e.g., accuracy, precision, recall) to quantify the model's predictive performance.

External Validation: Collaborating with external research institutions for independent validation studies to ensure the model's effectiveness across diverse populations.

6. Ethical Considerations and Patient Privacy: Informed Consent: Implementing processes to obtain informed consent from patients regarding the use of their data for predictive analytics.

Privacy Protection: Adhering to strict data security measures and compliance with legal and ethical standards to safeguard patient privacy.

7. Implementation and Iterative Feedback: Pilot Programs: Launching pilot programs in select healthcare facilities to gauge real-world performance and obtain feedback from healthcare professionals.

Iterative Adjustments: Making continuous adjustments to the methodology based on practical experiences and evolving research findings.

8. Education and Training: Healthcare education: Conduct training to familiarize healthcare professionals with methods and changes to predictive healthcare.
The approach is characterized by integration and reproducibility designed to increase the accuracy and validity of disease prediction based on past medical history while respecting the importance of ethics and patient privacy.

3.5.1 Software Required

- Anaconda Jupyter Notebook and Colab
- Machine Learning Algorithms
- Data Processing and Refining Algorithms
- Data Visualization Algorithms
- Django Framework

3.5.2 Hardware Required

- Memory and disk space required per user: 1GB RAM + 1GB of disk + .5 CPU core.
- Server overhead: 2-4GB or 10% system overhead (whatever is larger), .5 CPU cores, 10GB disk space.
- Port requirements: Port 8000 plus 5 unique, random ports per notebook.

3.5.3. Description of Solution Implemented

The disease prediction system implemented utilizes historical data to forecast the likelihood of an individual developing a specific medical condition. This comprehensive solution comprises data collection, preprocessing, and analysis, primarily employing Python, Pandas, and Scikit-Learn for data manipulation and feature extraction. It integrates data visualization tools like Matplotlib to gain insights from the data. Machine learning models, often built using TensorFlow or PyTorch, play a central role in predicting diseases. These models are trained on past data, taking into account various features and patterns to make accurate forecasts. Jupyter Notebooks facilitate model development and evaluation.
To ensure data privacy and security, the solution incorporates stringent practices and complies with healthcare data regulations. It may use Flask or Django for deploying the model as a web-based application, enabling easy access for both healthcare professionals and individuals—collaboration tools such as Scrum in team coordination. Continuous model improvement and adaptation are guided by performance metrics and cross-validation techniques. The system maintains version control using Git to track code changes effectively. Overall, this solution combines advanced technology, data science, and domain expertise to provide a valuable tool for disease prediction, ultimately enhancing early diagnosis and intervention for improved healthcare outcomes.
Implementations-

- Setup Anaconda Jupyter Notebook and Colab
- Configuring Google Drive and GitHub for Data Repository.
- Configuring ngrok pipelining
- Configuring interfaces with the model.
- Deploy the model on the cloud

3.5.4. Utility of the project

The disease prediction project holds immense utility in the realm of healthcare and has the potential to revolutionize the way we approach disease prevention and early intervention. First and foremost, this project empowers healthcare professionals with a valuable tool to make more informed decisions. By leveraging historical data and advanced machine learning models, the system can provide early warnings and predictions about the likelihood of specific diseases in patients. This enables healthcare providers to initiate timely interventions and treatments, potentially preventing the progression of diseases and improving patient outcomes.
Moreover, it enhances patient-eccentric care. Individuals can access personalized health risk assessments, allowing them to take proactive measures to mitigate risks. This not only promotes health awareness but also encourages patients to adopt healthier lifestyles and adhere to prescribed treatments.
On a broader scale, the project contributes to the optimization of healthcare resources. Identifying high-risk individuals and tailoring interventions accordingly, reduces unnecessary healthcare costs and hospitalizations. It also aids in resource allocation and planning for healthcare facilities and authorities, helping them prepare for the future healthcare needs of the population.
Additionally, this project can be precious in disease surveillance and epidemiology. It can identify disease outbreaks, track their spread, and provide critical data for public health agencies to formulate strategies to control and manage epidemics.
Overall, the disease prediction project is a powerful tool that not only improves individual health but also has far-reaching benefits for healthcare systems, resource management, and public health, ultimately contributing to a healthier and more informed society.

LITERATURE SURVEY

1. " Application of Machine Learning in Disease Prediction"[1]
The paper highlights the increasing use of machine learning in medical diagnosis, particularly in disease detection, owing to advancements in classification and recognition systems. These systems offer valuable data for aiding healthcare professionals in the early detection of life-threatening conditions, ultimately improving patient survival rates. The paper focuses on applying various classification algorithms to 3 different disease databases. Feature selection uses a backward modeling approach with the p-value test for each dataset. The study's findings underscore the potential of machine learning in facilitating the early detection of diseases, thus underscoring its significance in the field of healthcare.

2 “Symptoms Based Disease Prediction Using Machine Learning Techniques” [2]
The paper discusses the rapidly evolving field of Computer-Aided Diagnosis (CAD) in medical analysis, emphasizing the critical need for accurate diagnostic applications due to the potentially severe consequences of diagnostic errors in medicine. It highlights the importance of machine learning (ML) in CAD, especially when dealing with complex objects like body organs that cannot be

accurately identified through simple equations. Pattern recognition, a key component of CAD, necessitates training from examples. In the biomedical domain, the combination of pattern detection and ML holds the promise of enhancing the reliability of disease detection and diagnostic decision-making while maintaining objectivity.

3. "HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System"[3]
Heart disease is the leading cause of death worldwide and can benefit from early diagnosis. Adoption of clinical decision-making system (CDSS) may be useful in timely identification of individuals at risk. This study presents a cardiovascular risk prediction model (HDPM) developed specifically for CDSS. The hybrid model combines Functionality-Based Applied Spatial Clustering of Noise (DBSCAN) for detection and removal, Hybrid Synthetic Minority Oversampling Technique Regularization Nearest Neighbors (SMOTE-ENN) and Ensemble-Based Processing (SMOTE-EN) for product evaluation training.) Diseases. Predictive Random Forest) and previous studies were evaluated. Research results show that the performance of the proposed model is 95.90% and 98.40% accurate on the Statlog and Cleveland datasets, respectively.

Additionally, the Heart Disease CDSS (HDCDSS) model was developed to help doctors and nurses diagnose heart disease based on the current patient population. This positive approach has the potential to prevent deaths resulting from late diagnosis of heart disease by allowing early intervention.

4.” Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction” [4]
This research article is designed to provide current knowledge findings from archives, particularly research on the use of data mining in medical research, focusing on predicting cardiovascular disease. This study involves conducting various experiments to compare the performance of predictive data mining techniques on the same data.
The main findings of the study show that decision trees outperform other methods in terms of accuracy, while Bayesian classification also demonstrates competitive accuracy. However, methods such as KNN, neural networks and classification-based classification do not work very well in this case.

5.” A Survey of Machine Learning Approaches Applied to Gene Expression Analysis for Cancer Prediction”
Machine learning techniques have become powerful tools in developing cancer prediction models using gene expression and mutation data. This collection provides a comprehensive review of the latest cancer research focusing on various types of cancer, including breast, lung, kidney, ovarian, liver, central nervous system, and gallbladder. These studies use gene expression data for survival prediction, tumor identification, and classification. This review also includes biomarker studies related to these cancers. Covers all aspects of machine learning in cancer research, including cancer classification, prediction, gene biomarker analysis, and analysis of microarray and RNA-Seq data.
This study provides an in-depth analysis of the challenges associated with current cancer prediction models and the tools used to evaluate the performance of these genetics in soft tissue and tissue. Additionally, this paper explores how analysis of biomarker gene expression patterns can help predict future cancer risk and facilitate personalized treatment.

CONCLUSION

The integration of predictive analytics and machine learning in disease predictions based on past medical history represents a transformative leap toward proactive and personalized healthcare. This innovative approach harnesses the wealth of information embedded in electronic health records, paving the way for early detection, personalized interventions, and a paradigm shift from reactive to preventive healthcare strategies.

By leveraging advanced algorithms, these predictive models analyze past medical histories to identify subtle patterns, risk factors, and early indicators of potential health issues. The ability to discern genetic predispositions, lifestyle influences, and common symptoms enables healthcare professionals to tailor interventions to the unique characteristics of each individual. This individualized approach not only enhances the precision of healthcare but also empowers patients to actively participate in their well-being.

Preventive care, a cornerstone of this predictive healthcare paradigm, becomes a reality as healthcare providers can anticipate and mitigate health risks before they escalate. Targeted screening programs, lifestyle modifications, and personalized treatment plans contribute to the overall efficiency of healthcare systems, reducing the burden of reactive care.

However, the successful implementation of these predictive models requires addressing challenges such as data quality, interoperability, and ethical considerations. Rigorous validation of models and ongoing efforts to safeguard patient privacy is essential to build trust in these technologies.

In essence, disease predictions based on past medical history represent a pivotal advancement in healthcare, offering the potential to revolutionize patient care, improve health outcomes, and optimize resource utilization. As these technologies continue to evolve, their seamless integration into healthcare systems worldwide can redefine the healthcare landscape, making it more proactive, individualized, and focused on preventive measures for a healthier future.

FUTURE SCOPE

The future of disease predictions based on past medical history holds immense promise and potential for transformative advancements in healthcare. Several key areas indicate the evolving scope and opportunities for further development in this field.

Enhanced Predictive Models: Future research will likely focus on refining and enhancing predictive models through the incorporation of more sophisticated machine-learning algorithms and artificial intelligence techniques. Continuous innovation will lead to models with higher accuracy and the ability to predict a broader range of diseases.

Integration of multi-omics data: Integration of multi-omics data (such as genomics, proteomics, and metabolomics) will provide a better understanding of an individual's health. This holistic approach will make accurate predictions by taking into account various biological factors.

Real-Time Monitoring and Wearable Technology: The integration of real-time monitoring through wearable devices and continuous health tracking will enable the incorporation of dynamic data into predictive models. This shift towards real-time analytics will enhance the ability to identify trends and changes in health status promptly.

Population Health Management: Disease predictions based on past medical history will play a crucial role in population health management. Governments and healthcare providers can leverage these models to implement targeted interventions, allocate resources efficiently, and design preventive healthcare programs for specific demographic groups.

Blockchain for Data Security: The use of blockchain technology to secure and manage health data will become increasingly prominent. Blockchain ensures data integrity, security, and privacy, addressing concerns related to the ethical handling of sensitive medical information.

Collaboration and Standardization: Collaborative efforts among healthcare institutions, researchers, and technology developers will be vital for establishing standardized protocols and interoperability. A cohesive and standardized approach will facilitate the seamless exchange of health data, fostering a more integrated healthcare ecosystem.

Patient-centered approach: In the future, more emphasis will be placed on the patient process and people's participation in treatment will be ensured. Predicting individual diseases will enable patients to make informed decisions about their lifestyle, treatment plans and prevention.

In summary, the future scope of predicting diseases based on past medical history is characterized by broad and continuous technologies. Better understanding of health characteristics and providing adaptations to prevent patient health. These advances have the potential to revolutionize healthcare, improve patient outcomes, and contribute to the overall health of the global population.

REFERENCES

[1] P. S. Kohli and S. Arora, "Application of Machine Learning in Disease Prediction," 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018, pp. 1-4, doi: 10.1109/CCAA.2018.8777449.

[2] Hamsagayathri, P., & Vigneshwaran, S. (2021, February). Symptoms based disease prediction using machine learning techniques. In 2021 Third international conference on intelligent communication technologies and virtual mobile networks (ICICV) (pp. 747-752). IEEE.

[3] N. L. Fitriyani, M. Syafrudin, G. Alfian and J. Rhee, "HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System," in IEEE Access, vol. 8, pp. 133034-133050, 2020, doi: 10.1109/ACCESS.2020.3010511.

[4] Soni, J., Ansari, U., Sharma, D., & Soni, S. (2011). Predictive data mining for medical diagnosis: An overview of heart disease prediction. International Journal of Computer Applications, 17(8), 43-48.

[5] M. Khalsan et al., "A Survey of Machine Learning Approaches Applied to Gene Expression Analysis for Cancer Prediction," in IEEE Access, vol. 10, pp. 27522-27534, 2022, doi: 10.1109/ACCESS.2022.3146312.