IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed(Group –I) Journal Volume 12, Iss 01, January 2023

An In-Depth Theoretical Exploration of Data Mining in the Healthcare Sector: Techniques, Challenges, and Future Directions

[1]Saurabh Shandilya, [2]Keshav Dev Gupta, [3]Jameel Ahmed Qureshi, [4]Archna Bhardwaj,

[5]Priyanka Sharma, [6]Dr. Vaibhav Kumar Pradhan

[1]Professor, Department of Advance Computing, Poornima College of Engineering, Jaipur, Rajasthan, India.

[2]Associate Professor, Department of Advance Computing, Poornima College of Engineering, Jaipur, Rajasthan, India.

[3]Associate Professor, Department of Computer Engineering, Poornima University, Jaipur India.

[4]Assistant Professor, Department of Advance Computing, Poornima College of Engineering, Jaipur, India.

[5] Assistant Professor, Department of Geography, University of Technology, Jaipur, India.

[6]Senior Manager IT Audit, AU Small Finance Bank.

saurabh.shandilya@poornima.org keshav.gupta@poornima.org jameel.qurashi@poornima.org

archana.bhardwaj1@poornima.org keshav.gupta@poornima.org
Dr.Vaibhavpradhan26@gmail.com

ABSTRACT :- Data mining is a complex process aimed at uncovering hidden patterns within vast datasets. As a rapidly expanding discipline, it integrates statistics, visualization techniques, machine learning, and other data handling and knowledge discovery methods to reveal underlying trends and relationships within data. With the proliferation of the internet and mobile technology, data volumes are growing at an unprecedented rate, making analytics a critical focus across industries, both in IT and beyond. In particular, the healthcare sector is seeing increased exploration of data mining and machine learning applications to address data-intensive challenges. The demand for data warehouses is rising in healthcare, driven by the growing incidence of diseases, which is closely related to population growth and evolving lifestyles.
Keywords: Data Mining, Healthcare, Machine Learning, Industry..

1. INTRODUCTION

The rise of the internet and continuous technological advancements have ushered in a new era of digital interconnectivity, linking individuals with their numerous electronic devices. Recent innovations, along with distributed computing, have enabled real-time data capture, spanning not only digital content and transactions but also the real-time connection of devices and appliances, commonly referred to as the Internet of Things (IoT). The IoT represents the growing extension of the internet as more physical objects, such as consumer electronics and tangible assets, become networked. This includes technologies like mobile payments, Near Field Communication (NFC), image and voice recognition, and embedded sensors. This interconnected landscape has led to the rise of context-aware computing (CAC), which leverages environmental data related to the user to

improve interaction quality and deliver context-driven services across various domains. The continuous growth and enhancement of the internet, along with the broad availability of communication technologies, are rapidly transforming the very fabric of modern life.

2. U TILIZATION OF DATA MINING IN THE HEALTHCARE INDUSTRY

In today's world, the healthcare industry increasingly relies on data mining due to the vast amounts of data being generated. What was once considered a valuable tool is now becoming essential for the sector. Data mining is proving instrumental in reducing fraud in medical insurance and is being applied across various healthcare settings, including hospitals and research institutions. A key factor in choosing a data mining technique for healthcare is determining the type of information to be extracted. Data mining applications address numerous medical challenges, such as: • Categorizing patients into distinct groups • Identifying correlations between diseases and symptoms • Detecting patients with recurring health problems • Recognizing frequent patients and understanding the nature of their conditions • Predicting clinical diagnoses • Supporting medical research

The massive volumes of clinical data exchanged in healthcare are difficult to manage using traditional analysis methods. These vast datasets demand a more sophisticated approach, which is where data mining steps in, addressing challenges faced by both medical professionals and patients. Data mining enables healthcare organizations to make informed decisions by analyzing medical and financial data, helping to alleviate financial burdens. Its growing popularity in healthcare is driven by its role in developing decision support systems and non-invasive diagnostic techniques for rapidly spreading diseases. Some diagnostic procedures, including laboratory tests, can be costly, invasive, and uncomfortable for patients. For example, a biopsy for cervical cancer diagnosis is often a painful and unpleasant procedure. To streamline this, algorithms like K-means clustering can be applied to analyze cervical cancer patients, potentially providing more accurate predictions compared to traditional medical evaluations.

3. S IGNIFICANCE OF DATA MINING IN THE HEALTHCARE INDUSTRY

Public health represents a cost-efficient healthcare system designed to meet the health needs of communities and populations, with funding sourced directly from governments or state commissions. It ensures the financial sustainability of a healthy society by emphasizing equitable healthcare practices. In contrast to focusing on individual patients with specific medical conditions, public healthcare providers focus on delivering comprehensive care to entire communities. This approach ensures that the overall health of society remains positive, creating a safer, healthier environment for all. Public healthcare enables people to access a wide range of medical services, such as treatment for sinus congestion, sore throats, or even a fractured bone, at significantly lower costs compared to private healthcare systems. These government-funded healthcare institutions, including hospitals and clinics, are conveniently located to serve the population effectively, making healthcare more accessible to everyone.

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed(Group –I) Journal Volume 12, Iss 01, January 2023

Public health places a strong emphasis on improving the health of entire populations rather than addressing individual cases. This focus on collective health requires collaboration and aims to tackle disease prevention, treatment, and care from a community-wide perspective. While public healthcare systems primarily focus on managing health at the community level, ensuring widespread access to medical care remains a critical challenge for many developing countries. This access is essential for maintaining a healthy population, as health disparities often arise from a lack of available healthcare services. In recent years, electronic health records (EHRs) have become the norm across many healthcare organizations. These records provide healthcare providers with better access to vast amounts of patient data, enabling them to improve both the efficiency and quality of their care.

With the wealth of data now available, healthcare organizations have increasingly turned to data mining to optimize their operations. Data mining, which has been widely used since the 1990s for applications like fraud detection in financial services, is now finding its place in healthcare as well. Today, many healthcare organizations across the globe are realizing the immense value that medical data analysis and predictive analytics can bring. Through data processing, the goal is often to uncover meaningful patterns and insights from large datasets. These insights can then be applied to predict future trends in both medical and business contexts, helping organizations to make more informed decisions about patient care and operational strategies.

Data mining in healthcare offers a host of benefits. It can significantly reduce operational costs, improve efficiencies, enhance patient quality of life, and, most importantly, contribute to saving more lives by making healthcare more proactive and data-driven. The technology investments that companies make toward these goals enable more effective data utilization, resulting in better decision-making and outcomes. Whether referred to as "analytics," "business intelligence," or "data processing," the central concept of data mining remains consistent: analyzing large datasets to recognize patterns and using those patterns to predict future events. In healthcare, this ability to predict and forecast trends has enormous potential, from improving patient outcomes to driving better financial decisions. Thus, data mining has become a crucial tool for healthcare organizations striving to enhance their service delivery and outcomes.

4. HEALTHCARE DATA MINING TECHNIQUES AND ALGORITHMS

The The rapid advancements in healthcare and medicine have led to an explosion of data generation, including electronic health records, clinical reports, and patient statistics. Despite this abundance of data, many healthcare systems struggle to harness the full potential of their data. Data mining emerges as a powerful tool to uncover hidden insights within these massive datasets, transforming the way healthcare is delivered and managed. By applying sophisticated algorithms and techniques, healthcare professionals can gain valuable insights that improve diagnosis, treatment, and patient outcomes. Here, we explore several key data mining techniques used in healthcare, illustrating their applications and benefits in this critical field.

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed(Group –I) Journal Volume 12, Iss 01, January 2023

1. Anomaly Detection

Anomaly detection is a crucial technique in data mining used to identify unusual patterns or outliers in large datasets. In the context of healthcare, detecting anomalies can be instrumental in identifying unexpected changes in patient health data or system performance. For example, anomaly detection methods such as support vector data description (SVDD), density-induced SVDD, and Gaussian mixture models are used to analyze liver disorder datasets. These methods help pinpoint abnormalities that could indicate serious health conditions or data entry errors. Anomaly detection techniques often employ metrics like the Area Under Curve (AUC) to evaluate their effectiveness. By identifying anomalies, healthcare providers can address potential issues early, leading to improved diagnostic accuracy and patient care.

2. Clustering

Clustering involves grouping data points into clusters based on shared characteristics, which helps in understanding the structure and patterns within the data. In healthcare, clustering is used to segment patients into groups with similar health conditions or treatment needs. Techniques such as vector quantization, which includes algorithms like K-means, X-means, and K-medoids, are commonly applied. For instance, clustering can be used to predict patient readmissions by analyzing clinical data and laboratory results. The quality of clustering is assessed using indices like the Davies-Bouldin Index, which measures the separation and compactness of clusters. Effective clustering allows healthcare professionals to identify high-risk patient groups, tailor interventions, and improve treatment outcomes.

3. Classification

Classification is a technique that assigns data points to predefined categories based on their attributes. In healthcare, classification algorithms are used to predict patient diagnoses or outcomes. For example, statistical methods such as Mahalanobis distance (MD) and Mahalanobis space (MS) are employed to differentiate between clusters and identify abnormal data points. The Mahalanobis Taguchi System (MTS) is used to develop predictive models for conditions like pressure ulcers. Classification techniques are particularly useful in addressing class imbalance issues, where some health conditions are rare compared to others. By accurately classifying patients, healthcare providers can make informed decisions about treatment and care, leading to better patient management.

4. Discriminant Analysis

Discriminant Analysis is a statistical technique used to classify data based on measurements from unlabeled observations. Linear Discriminant Analysis (LDA) is a common method that estimates the likelihood of a data point belonging to a particular class. LDA has been used in research to predict the severity of diseases such as Parkinson's by analyzing both motor and non-motor symptoms. This method helps in understanding the relationships between different health indicators and provides insights into disease progression. By incorporating statistical relationships among

predictor variables, discriminant analysis offers valuable information for making accurate diagnoses and treatment decisions.

5. Decision Trees

Decision trees are a versatile tool used in healthcare to make predictive decisions based on data analysis. A decision tree is a flowchart-like structure where each branch represents a decision rule, and each leaf node represents an outcome. Decision trees help in classifying patients, predicting outcomes, and assessing treatment options. Various decision tree algorithms, including the Gini Index, Information Gain, and Gain Ratio, are used to evaluate the effectiveness of different branches. Decision trees can also be pruned to reduce complexity and improve performance. By providing clear decision rules, decision trees enable healthcare professionals to make data-driven decisions and improve patient care.

6. Neural Networks

Neural networks are inspired by the structure of the human brain and are used to analyze complex datasets in healthcare. They consist of interconnected nodes, or neurons, organized in layers. Neural networks are particularly effective for tasks such as image recognition, genomic data analysis, and electronic health record processing. For example, neural networks can be used to detect tumors in medical images, predict patient outcomes, and analyze clinical notes. Their ability to learn from large amounts of data and identify patterns makes them a valuable tool for improving diagnostic accuracy and developing personalized treatment plans.

7. Support Vector Machines (SVM)

Support Vector Machines (SVMs) are a robust and sophisticated technique used extensively for both classification and regression tasks. The core functionality of SVMs involves identifying the optimal hyperplane that effectively separates different classes within a given dataset. This hyperplane is strategically positioned to maximize the margin between the classes, ensuring the best possible separation.

In the context of healthcare, SVMs play a crucial role in various applications. They are employed to classify diseases based on patient data, which is instrumental in distinguishing between different medical conditions. For example, SVMs can be used to differentiate between benign and malignant tumors by analyzing features extracted from medical imaging. Additionally, they are valuable in predicting the likelihood of certain health conditions based on patient records and test results.

One of the significant advantages of SVMs is their ability to handle high-dimensional data effectively. Healthcare datasets often contain numerous variables, and SVMs are particularly adept at managing these complexities. By utilizing SVMs, healthcare professionals can uncover intricate relationships between different variables that might not be apparent through simpler methods.

The application of SVMs in healthcare enables more precise predictions and enhances diagnostic capabilities. For instance, by training SVM models on historical patient data, healthcare providers

can develop predictive tools that assist in early diagnosis and personalized treatment planning. This results in improved patient outcomes and more efficient use of medical resources.

.

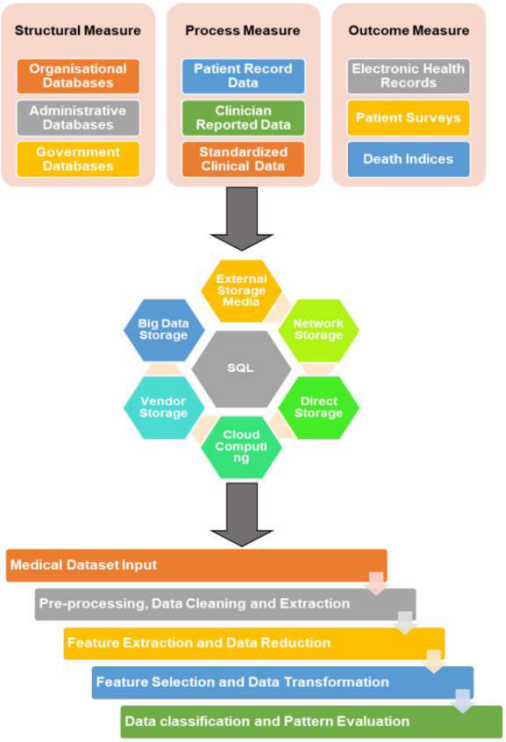


Figure 1 Implementation of ID3 algorithm on patient data

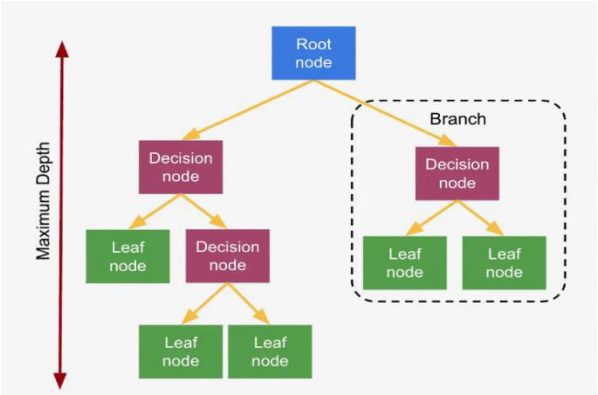


Figure 1.1 Implementation of ID3 algorithm on patient data

Example Dataset

Let's assume we have a dataset of patients with the following attributes:

1. **Age**: Young, Middle-aged, Old
2. **Gender**: Male, Female
3. **Symptoms**: Fever, Cough, Headache
4. **Diagnosis**: Flu, Cold, Allergy

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed(Group –I) Journal Volume 12, Iss 01, January 2023

Step-by-Step Implementation

1. Data Preparation

Start with a dataset of patient records with attributes and their corresponding diagnoses.

Age	Gender	Symptoms	Diagnosis
Young	Male	Fever	Flu
Middle-aged	Female	Cough	Cold
Old	Male	Fever	Flu
Young	Female	Cough	Cold
Old	Female	Headache	Allergy
Middle-aged	Male	Headache	Allergy

2. Calculate Entropy of the Dataset

Calculate the entropy of the entire dataset to measure the impurity.

$$\text{Entropy}(S) = -\sum_{i=1}^{c} p_i \log_2(p_i)$$

where pip_ipi is the probability of each class.

- Probability of Flu = 2/6
- Probability of Cold = 2/6
- Probability of Allergy = 2/6

$$\text{Entropy}(S) = -\left(\frac{2}{6}\log_2\frac{2}{6} + \frac{2}{6}\log_2\frac{2}{6} + \frac{2}{6}\log_2\frac{2}{6}\right) \approx 1.585$$

3. Calculate Information Gain for Each Attribute

For Age:

- **Young**: 2 Flu, 2 Cold (Entropy = 1.0)
- **Middle-aged**: 1 Cold, 1 Allergy (Entropy = 1.0)
- **Old**: 1 Flu, 1 Allergy (Entropy = 1.0)

Calculate weighted average entropy for Age:

$$\text{Entropy}_{\text{Age}} = \frac{4}{6} \cdot 1.0 + \frac{2}{6} \cdot 1.0 + \frac{2}{6} \cdot 1.0 = 1.0$$

$$\text{Information Gain}_{\text{Age}} = \text{Entropy}(S) - \text{Entropy}_{\text{Age}} = 1.585 - 1.0 = 0.585$$

For Gender:

- **Male**: 2 Flu, 1 Allergy (Entropy = 0.918)
- **Female**: 2 Cold, 1 Allergy (Entropy = 0.918)

Calculate weighted average entropy for Gender:

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed(Group –I) Journal Volume 12, Iss 01, January 2023

$$\text{Entropy}_{\text{Gender}} = \frac{3}{6} \cdot 0.918 + \frac{3}{6} \cdot 0.918 = 0.918$$

$$\text{Information Gain}_{\text{Gender}} = \text{Entropy}(S) - \text{Entropy}_{\text{Gender}} = 1.585 - 0.918 = 0.667$$

For Symptoms:

- **Fever**: 2 Flu (Entropy = 0.0)
- **Cough**: 2 Cold (Entropy = 0.0)
- **Headache**: 2 Allergy (Entropy = 0.0)

Calculate weighted average entropy for Symptoms:

$$\text{Entropy}_{\text{Symptoms}} = \frac{2}{6} \cdot 0.0 + \frac{2}{6} \cdot 0.0 + \frac{2}{6} \cdot 0.0 = 0.0$$

$$\text{Information Gain}_{\text{Symptoms}} = \text{Entropy}(S) - \text{Entropy}_{\text{Symptoms}} = 1.585 - 0.0 = 1.585$$

4. Select the Best Attribute

The attribute with the highest information gain is "Symptoms" (1.585). So, we choose "Symptoms" as the root node.

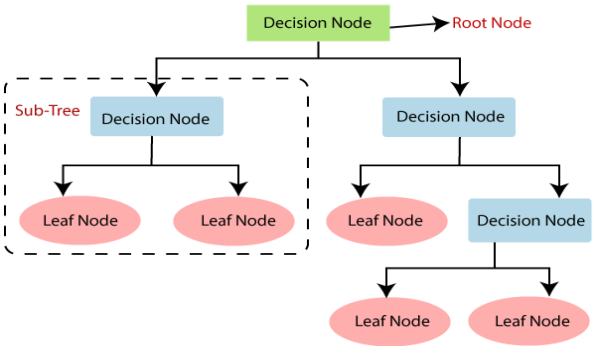
5. Create Subsets Based on the Best Attribute

Split the dataset into subsets based on "Symptoms":

- **Fever**: Diagnoses = Flu
- **Cough**: Diagnoses = Cold
- **Headache**: Diagnoses = Allergy

Each subset now has a single class, so no further splitting is needed.

6. Construct the Decision Tree



Swarm Intelligence

Swarm Intelligence, particularly through the Particle Swarm Optimization (PSO) technique, is a powerful method for finding optimal or near-optimal solutions within extensive search spaces. The essence of PSO lies in its ability to explore and exploit large areas of the solution space effectively.

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed(Group –I) Journal Volume 12, Iss 01, January 2023

By optimizing the parameters used in classification models, PSO enhances the overall classification performance. The approach is particularly advantageous when dealing with complex datasets, as it refines the search for the best parameters, leading to more accurate and efficient classification results.

K-Nearest Neighbor (K-NN)

The K-Nearest Neighbor (K-NN) algorithm is a widely-used technique for classifying data based on its similarity to other data points. This instance-based approach classifies a new data point by comparing it to its nearest neighbors in the feature space. Each sample in the training dataset is considered during the classification of new instances. The main advantage of K-NN is its simplicity and the fact that it retains all information from the training data, which means it doesn't require a model to be explicitly built. However, this approach can be computationally intensive, especially with large datasets, as each classification involves comparing the new sample to every point in the training set. Despite this, K-NN is valuable in medical diagnostics due to its ability to classify complex patterns based on historical data.

Logistic Regression

Logistic Regression (LR) is a statistical method used for binary classification and can also handle continuous and discrete outcomes. It involves calculating a linear combination of input features and passing this combination through a logistic function to produce probabilities. This method is favored for its simplicity and interpretability, making it a common choice for various classification tasks. LR is especially useful in healthcare for predicting the probability of a certain medical condition or outcome based on patient features, and its results are often comparable to more complex models while being easier to implement and understand.

Bayesian Classifier

The Bayesian Classifier is a robust technique that excels in handling missing data and provides high computational efficiency. It operates based on Bayes' Theorem, which calculates the probability of a sample belonging to a particular class. This classifier is particularly effective in situations where the data is incomplete or noisy. The Naive Bayes variant, which assumes independence among features, is widely used due to its simplicity and effectiveness. By leveraging statistical probabilities, Bayesian classifiers offer a clear understanding of the likelihood of various outcomes, enhancing predictive accuracy. They are commonly employed in medical settings to estimate the probability of disease based on patient attributes and historical data. The implementation of the Naive Bayes algorithm can be visualized in Figure 2, which illustrates the process of predicting class membership based on available features.

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed(Group –I) Journal Volume 12, Iss 01, January 2023

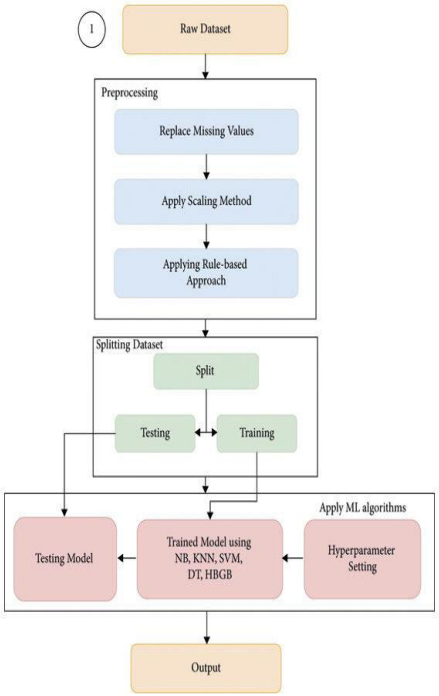


Figure 2 The Naive Bayes method was applied to the patient data

Support Vector

As a result of the fact that Support Vector Machines have proven to be quite effective in a wide variety of pattern categorization tasks, they have recently attracted a significant amount of interest. The Support Vector Machine, or SVM, is a form of supervised machine learning. The SVM algorithm accurately predicts the occurrence of disease and optimally classifies the classes by constructing the margin among two data clusters. This is accomplished by projecting the disease predicting characteristics in a multidimensional hyper plane. This approach is able to reach great precision because it makes use of kernels, which are nonlinear features. It has been established that the support vector approach (SVM), which has excellent generalisation efficiency, is advantageous in the management of classification responsibilities. The method works on lowering the upper boundary of the generalisation error by using a model that is geared toward reducing the amount of risk that the structure poses. Training a support vector machine is equivalent to finding a solution to a quadratic programme problem that is linearly constrained. The method is frequently applied in the field of medical diagnosis. In the SVM technique, the two variables that have the potential to regulate generalisation are the training mistake and the capacity of the assessed learning machine. Both of these variables are considered to be capacities. The training error rate can be adjusted by making changes to the attributes that are contained inside the classifiers. The precision of the data mining approach varies between the practise sets and the test sets according to the characteristics of the information sets as well as the size of the data sets. When it comes to healthcare data sets, highly unbalanced data sets are one of the most common aspects. This means that the majority classifier and the minority classifier are not in a balanced state, which leads to an inaccurate forecast when the classifiers are run. In addition to these characteristics, information sets for healthcare often contain values that are missing. Because there is typically such a little amount of information that can be accessed, the sample size of the information is frequently considered to be

another attribute. There is not an effective method of information mining that can solve all of these issues at once.

- **Neural Network**

It is commonly known that neural networks can generate results in practical applications that are of an incredibly exact nature. The neural network was trained with the database through the utilisation of the feed-forward neural network model, variable learning speed, and the momentum learning algorithm back-propagation. Following is an explanation of the construction of the model: it starts with the input of clinical information and then moves on to the development of an ANN algorithm. Following the completion of the training model, it is able to generate prediction results.

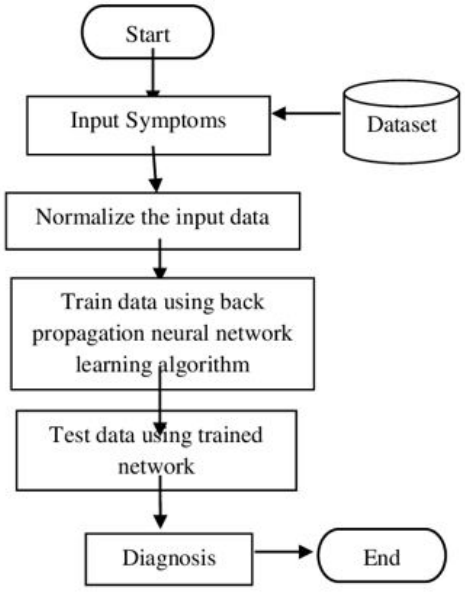


Figure Implementation of neural network algorithm on clinical data.

The random division of clinical data into two equal halves is where the computational stages of an algorithm for a neural network get started. The one that is being used for testing is different from the one that is being used for practise. Randomization is used to determine the initial weight that is given to each function. The mistakes that were calculated are employed in order to modify the weight of all of the qualities. When the mistakes satisfy the termination criteria, the final weight of each feature is determined. The procedure is carried out on a number of different occasions. After constructing the training models, we are able to derive the performance outcomes from the test information and calculate them. Figure 1.6 depicts the execution of the neural network method with regard to clinical information.

- **Hybrid**

One of the most significant challenges that the healthcare industry is currently experiencing is disease prediction. In order to diagnose diseases, scientists employ a wide variety of information mining approaches. This is motivated by the rising death rate of patients worldwide. Each approach has a set of advantages as well as some drawbacks. Every algorithm that is utilised by every approach incorporates several characteristics that are helpful in the process of diagnosing the illness. The term "hybridization" will be used to refer to the output mixture here. In the process of disease diagnosis, the application of hybrid data mining approaches can produce some encouraging outcomes. The methodology behind the hybrid data mining technique is illustrated in figure 1.7.

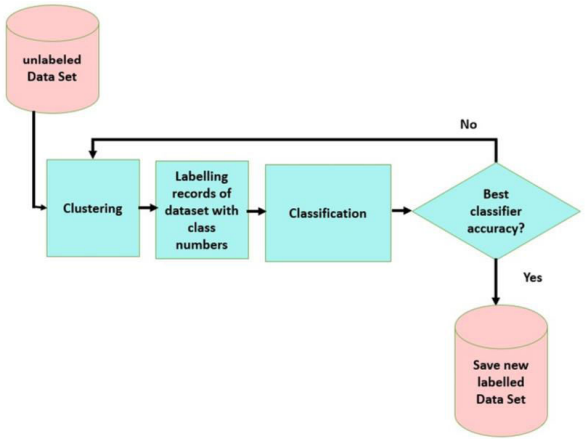


Figure 1.7 Approach of hybrid data mining technique.

5. CONCLUSION

In the field of medicine, achieving a high level of accuracy is crucial, and recent advancements in data mining techniques have shown that these methods can reach the desired precision. Data mining tools have become increasingly advanced, enabling the extraction of unique and actionable patterns from extensive datasets. This is particularly valuable in complex areas such as carcinogenesis, the process through which normal cells transform into cancerous cells. Carcinogenesis presents a significant challenge to researchers, as controlling and understanding this process involves intricate and voluminous data that traditional methods struggle to handle. Specialized data mining tools, when applied to Electronic Data Warehouses (EDWs), can unravel these complexities by utilizing sophisticated algorithms to extract meaningful insights. This approach enhances the accuracy and depth of cancer research by revealing patterns and relationships that were previously obscured within large datasets.

The integration of data mining techniques with data stored in EDWs offers a comprehensive approach to analyzing and understanding complex medical conditions. Techniques such as anomaly detection, clustering, and classification provide critical insights into disease progression and patient characteristics. For example, clustering algorithms can group similar patient data to identify patterns related to disease development, while classification algorithms can predict disease likelihood based on various factors. This application of data mining not only improves the understanding of conditions like carcinogenesis but also aids in developing more effective treatment strategies and preventive measures. By leveraging these advanced techniques, researchers and clinicians can make more informed decisions, ultimately leading to better patient outcomes and advancements in medical science.

REFERENCES

1. Hansapani Rodrigo (2021) ,” Exploratory Data Mining Techniques (Decision Tree Models) for Examining the Impact of Internet-Based Cognitive Behavioral Therapy for Tinnitus: Machine Learning Approach”,
2. AntonellaGuzzo (2021),” Process mining applications in the healthcare domain: A comprehensive review”,
3. V Klochko et al (2020),”Data mining of the healthcare system based on the machine learning model developed in the Microsoft azure machine learning studio”,

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed(Group –I) Journal Volume 12, Iss 01, January 2023

4. Adam Bohr et al (2020),"The rise of artificial intelligence in healthcare applications", Artificial
5. Suneeta S. Raykar (2020)," Cognitive Analysis of Data Mining Tools Application in Health Care Services",
6. S. Dhanalakshmi(2020) ," Survey on Information Mining Procedures Utilized in Healthcare Services"
7. Adam Bohr(2020)," The rise of artificial intelligence in healthcare applications",
8. Sayantan Khanra (2020)," Big data analytics in healthcare: a systematic literature review",
9. Hui Yang (2020)," The Use of Data Mining Methods for the Prediction of Dementia: Evidence from the English Longitudinal Study of Aging",
10. Dharmpal Singh (2019)," Cognitive Social Mining Analysis Using Data Mining Techniques",
11. S. Aarathi (2019),"Impact of Healthcare Predictions with Big Data Analytics and Cognitive Computing Techniques", International Journal of Recent Technology and Engineering