DISEASE PREDICTION USING MACHINE LEARNING, DEEP LEARNING AND DATA ANALYTICS

Editors:

Geeta Rani

Vijaypal Singh Dhaka

Pradeep Kumar Tiwari

Bentham Books

Disease Prediction using Machine Learning, Deep Learning and Data Analytics

Edited by

Geeta Rani
Department of Computer and Communication Engineering
Manipal University Jaipur
Jaipur, India

Vijaypal Singh Dhaka
Department of Computer and Communication Engineering
Manipal University Jaipur
Jaipur, India

&

Pradeep Kumar Tiwari
Dr. Vishwanath Karad MIT
World Peace University
Pune, India

Disease Prediction using Machine Learning, Deep Learning and Data Analytics

Editors: Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari

ISBN (Online): 978-981-5179-12-5

ISBN (Print): 978-981-5179-13-2

ISBN (Paperback): 978-981-5179-14-9

First published in 2024.

Bentham Science Publishers Pte. Ltd.
80 Robinson Road #02-00
Singapore 068898
Singapore
Email: subscriptions@benthamscience.net

CONTENTS

FOREWORD

It is my pleasure to write the foreword for the book titled **“Disease Prediction using Machine Learning, Deep Learning and Data Analytics”**. The book covers the role of machine learning in boosting the immunity of a person, role of federated learning in healthcare, role of data mining in developing a medical support system, and role of AI in establishing interaction between human brain and computer.

This book covers the detection of diabetic retinopathy using machine learning algorithms, deep learning based model for conversion of 2-D images into to 3-D images, developing a decision support system for prediction of asthma attack, early prediction of eye diseases, computer-aided bio-medical tools for disease identification, deep learning based systems for medical data classification and AI-based chatbot system for healthcare industry.

The book “Disease Prediction using Machine Learning, Deep Learning and Data Analytics”, gives a clear idea about the deep learning techniques employed for analysis and classification of imagery data. The data mining algorithms for knowledge extraction and feature extraction techniques attract the readers working in the field of medical image analysis. The book provides the mechanisms involved in designing and developing the clinical decision support systems.

“Disease Prediction using Machine Learning, Deep Learning and Data Analytics”, is a must read book for the academicians, researchers and students working in the field of applications of machine learning and deep learning for disease diagnosis and prognosis. The book is important to read for the clinical experts who are keen to adopt the techno-tools as assistants for diagnosis and prognosis of diseases.

I would like to congratulate the Editor in Chief, Dr. Geeta Rani and Associate Editors, Dr. Pradeep Kumar Tiwari and Dr. Vijaypal Singh Dhaka for bringing the ideas of academicians and research community together at a single platform. I strongly believe that their expertise in the field of machine learning, cloud computing and medical image analysis will be effective in attracting the readers in the field.

Dharm Singh Jat
Namibia University of Science & Technology
Windhoek, Namibia

PREFACE

In the recent era, the use of data analytics and machine learning algorithms has been observed in the arena of the medical field. Literature shows the successful application of data analytics and machine learning techniques for making predictions using real-time data collected from medical fields. The efficacy of machine learning models in image processing, big data analytics, object detection, automatic extraction, and tailoring of features is a great motivation for employing these models in the medical field. A boom in the use of machine learning and deep learning models is observed since the last decade. These models automatically extract the features from medical images, identify the most prominent features and predict diseases such as pneumonia, COVID-19, emphysema, lung tuberculosis, tumor, *etc.* can be predicted by training the deep learning model with chest radiographs and CT scans. These models not only predict the disease but are also useful in visualizing the infection in the organs. For reliable prediction, there is a need to design the custom architecture of the model. The architecture designer must focus on the size of the dataset, versatility, and quality of the dataset, types and number of predictions to be provided. The architecture is also dependent on the type of analysis required for disease prediction.

Literature reveals a lot of information about the design of methods for disease prediction.

But, poor availability of systematic information at one source becomes challenging for the students, academicians as well as researchers working in this field. Researchers face problems in identifying suitable algorithms for pre-processing, transformations, and integration of clinical data. They also seek different ways to build models, and prepare data sets for training and evaluating the models. Moreover, it becomes significant for them, to observe the impact of decision-making strategies on the accuracy and precision of the predictive models designed on the basis of techniques such as Logistic Regression, Neural Networks, Decision Trees, and Nearest Neighbors. Thus, there is a strong need of providing well-organized study material with practical aspects and validation. The book smartly fills the gaps.

This book invited ideas, proposals, review articles and experimental works from the researchers working in the field. The systematic organization of the research works in the field of applying machine learning for disease prediction will be fruitful in providing insights to readers about the existing works and the gaps available in the field. This book is a significant contribution towards providing a detailed study of data analytics algorithms and machine learning techniques for disease prediction. The book includes a rigorous review of related literature, methodology for data set preparation, model building, training, and testing the model. It contains a comparative analysis of versatile algorithms applied for making predictions in the challenging arena of medical science and disease prediction. The provides good insight into the topics such as Data Analytics, Machine Learning, Deep Learning, Information Retrieval from medical data, Data Integration, Prediction Models, Medical Data Analysis, Medical Decision Support systems, Federated Learning in Healthcare, and Medical Image Reconstruction.

The book is a companion and a must-read, for academicians, people from industries, graduate and post-graduate students, researchers, physicians and for everyone who is involved in the fields of medicine, data analytics or machine learning directly or indirectly. The book is compiled in such a way that each chapter is sufficient to give a complete study set from problem formulation to its solutions. All chapters are independent of each other and can be studied individually without consulting other chapters.

Each chapter starts with an abstract, important key terms, and an introduction to the topic. It is followed by related works, challenges identified, methodology, and experimental results. The chapter ends with the concluding remarks and future directions.

This book includes chapters in the following research areas:

- Review in the fields of Data Analytics, Machine Learning, and Medical Data Analysis.
- Federated Learning in Healthcare.
- 3-Dimensional Image Reconstruction.
- Applications and Practical Systems for Healthcare.
- Information Retrieval from medical data.
- Data Integration.
- Prediction Models.
- Clinical Decision Support Systems.
- Computer-Aided Diagnosis.
- Mobile Imaging for Biomedical Applications.

A brief summary of book chapters is given below:

Chapter 1: Role of Federated Learning in Healthcare: A Review

In this chapter, the authors provide a detailed comparative study of the different deep learning-based models employed in federated learning. They discussed how efficiently the model can classify chest radiographs into Covid-19, pneumonia, and normal categories. This chapter provides the benchmarking information and analysis for the researchers looking forward to developing deep learning-based applications of federated learning in healthcare.

Chapter 2: Role of Artificial Intelligence in 3-D Bone Image Reconstruction: A Review

This chapter presents a review of the bone imaging techniques and techniques applied for the conversion of two-dimensional images into three-dimensional form. It also gives directions for developing patient-specific and organ-specific optimized techniques for 3-D reconstruction.

Chapter 3: Role of Machine Learning and Deep Learning Techniques in Detection of Disease Severity: A Survey

This chapter explores the role of machine learning and deep learning techniques in the detection of disease severity. It presents a survey of the latest methodologies and algorithms employed in analyzing medical data to predict and assess the severity of various diseases, empowering clinicians with valuable insights for personalized treatment plans. The chapter highlights the advantages and drawbacks of different ML and DL techniques employed for prediction of disease severity.

Chapter 4: Computer-Aided Bio-Medical Tools for Disease Identification

This chapter investigates computer-aided biomedical tools for disease identification. It discusses the development and utilization of innovative software tools and techniques that assist in the identification and diagnosis of diseases, augmenting healthcare professionals' decision-making process. The chapter highlights the importance of computer-aided bio-medical tools as techno-assistants for health experts.

Chapter 5: Prognosis of Dementia using Machine Learning

In this chapter, the authors discuss the prognosis of dementia using machine learning. They explore the potential of machine learning algorithms in predicting the progression and prognosis of dementia, offering valuable insights for early interventions and personalized care plans.

Chapter 6: A Clinical Decision Support System for Effective Identification of Onset of Asthma Disease

This chapter presents a clinical decision support system for the identification of asthmatics in two different cohorts representing rural and urban populations in India. It provides details about developing a hybrid decision support system by uniquely combining unsupervised and supervised learning techniques.

Chapter 7: Applying Deep Learning and Computer Vision for Early Diagnosis of Eye Diseases

This chapter presents a study to raise awareness about various eye disorders. It provides a discussion on the role of computer vision, image processing, and deep learning techniques in the early diagnosis of the disease. Thus, it may prove useful for enhancing early disease treatment and minimizing the chances of blindness.

Chapter 8: The Fusion of Human-Computer Interaction and Artificial Intelligence Leads to the Emergence of Brain Computer Interaction

In this chapter, the authors discuss the Components Brain Computer Interface, its characteristics and challenges. They provide details of how conventional classifiers are replaced with Convolutional Neural Networks (CNNs). The chapter also reveals that the EEG signals from the brain can be linked seamlessly to mechanical systems *via* BCI applications, making it a rapidly growing technology. The presented technology has applications in the fields such as Artificial Intelligence and Computational Intelligence.

Chapter 9: Mining Standardized EHR Data: Exploration, Issues, and Solution

This chapter focuses on mining standardized Electronic Health Records (EHR) data, providing an in-depth exploration. This chapter examines the process of extracting knowledge and insights from standardized EHR data through data mining techniques. It explores the challenges and opportunities associated with mining EHR data, including data quality issues, data integration challenges, and ethical considerations of handling sensitive patient information. Additionally, the chapter presents innovative solutions and methodologies for effectively mining EHR data to support various healthcare applications such as clinical decision-making, predictive analytics, and population health management.

Chapter 10: Role of Database in Epidemiological Situation

This chapter presents the crucial role of databases in epidemiological situations. It highlights the significance of databases in epidemiological research, providing a comprehensive overview of their role in data collection, management, and analysis. In this chapter, the authors explore different types of databases commonly used in epidemiology, including disease surveillance systems. Moreover, the chapter discusses the challenges and considerations associated with database implementation, such as data standardization, and privacy protection.

Geeta Rani
Department of Computer and Communication Engineering
Manipal University Jaipur
Jaipur, India

Vijaypal Singh Dhaka
Department of Computer and Communication Engineering
Manipal University Jaipur
Jaipur, India

&

Pradeep Kumar Tiwari
Dr. Vishwanath Karad MIT
World Peace University
Pune, India

INTRODUCTION

The healthcare industry is undergoing a transformative phase, driven by technological advancements and data-driven solutions. The integration of cutting-edge technologies, such as Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL), have opened new horizons for improved healthcare delivery, disease diagnosis, and patient care. Additionally, the utilization of computer-aided tools helps clinicians and researchers to make more accurate and timely decisions. This motivated us to provide a concise information regarding technology-based solutions in healthcare.

This book aims to explore and discuss the significant role of technology and data in various domains of healthcare. It delves into the latest research and developments in the field, providing comprehensive reviews, surveys, and case studies on topics ranging from federated learning to disease prognosis and biomedical tools. It also delves into two important aspects of leveraging technology and data in healthcare, mining standardized EHR data and the role of databases in epidemiological situations.

Each chapter focuses on a specific aspect, highlighting the potential impact of technology on enhancing healthcare practices and outcomes. The brief outline of each chapter is given below:

Chapter 1 examines the role of federated learning, a distributed machine learning approach that enables collaborative analysis of sensitive healthcare data while preserving privacy and security. It presents a comprehensive review of the applications and benefits of federated learning in healthcare settings.

In Chapter 2, the focus shifts to the role of artificial intelligence in a 3-D bone image reconstruction. This chapter provides an in-depth review of the advancements in AI techniques for reconstructing detailed and accurate three-dimensional bone images, aiding in better diagnosis and treatment planning.

Chapter 3 explores the role of machine learning and deep learning techniques in the detection of disease severity. It presents a survey of the latest methodologies and algorithms employed in analyzing medical data to predict and assess the severity of various diseases, empowering clinicians with valuable insights for personalized treatment plans.

Chapter 4 investigates computer-aided biomedical tools for disease identification. It discusses the development and utilization of innovative software tools and techniques that assist in the identification and diagnosis of diseases, augmenting healthcare professionals' decision-making process.

In Chapter 5, the authors discuss the prognosis of dementia using machine learning. This chapter explores the potential of machine learning algorithms in predicting the progression and prognosis of dementia, offering valuable insights for early interventions and personalized care plans.

Chapter 6 introduces a clinical decision support system for the effective identification of the onset of asthma disease. It explores how advanced technologies, including machine learning and data analytics, can be integrated into clinical workflows to enhance the early detection and management of asthma.

Chapter 7 delves into the application of deep learning and computer vision for early diagnosis of eye diseases. It discusses the utilization of these technologies to analyze medical images, enabling early detection and intervention in conditions such as diabetic retinopathy and glaucoma.

Chapter 8 explores the fusion of human-computer interaction and artificial intelligence, leading to the emergence of brain-computer interaction. It delves into the advancements in this interdisciplinary field, highlighting its potential to revolutionize healthcare through direct communication between the brain and computer.

Chapter 9 focuses on mining standardized Electronic Health Records (EHR) data, providing an in-depth exploration. This chapter examines the process of extracting knowledge and insights from standardized EHR data through data mining techniques. It explores the challenges and opportunities associated with mining EHR data, including data quality issues, data integration challenges, and ethical considerations of handling sensitive patient information. Additionally, the chapter presents innovative solutions and methodologies for effectively mining EHR data to support various healthcare applications such as clinical decision-making, predictive analytics, and population health management.

Chapter 10 delves into the crucial role of databases in epidemiological situations. This chapter highlights the significance of databases in epidemiological research, providing a comprehensive overview of their role in data collection, management, and analysis. It explores different types of databases commonly used in epidemiology, including disease surveillance systems. Moreover, the chapter discusses the challenges and considerations associated with database implementation, such as data standardization, and privacy protection.

In this book, we aim to discuss the advancements and potential of technology and data-driven approaches in healthcare.

DEDICATION

This book is dedicated to all authors who contributed their valuable work in this book. Further, the book is dedicated to everyone who supported the editors Dr. Geeta Rani, Dr. Pradeep Kumar Tiwari and Dr. Vijaypal Singh Dhaka directly or indirectly in completion of this book well in time.

List of Contributors

Anu Saini	GB Pant DSEU, Okhla-1 Campus, Delhi, India
E. Francy Irudaya Rani	Department of Electronics and Communication Engineering, Francis Xavier Engineering College, Tirunelveli, Tamil Nadu, India
E. Fantin Irudaya Raj	Department of Electrical and Electronics Engineering, Dr. Sivanthi Aditanar College of Engineering, Thoothukudi, Tamil Nadu, India
Geeta Rani	Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India
Heta Patel	Manipal University Jaipur, Jaipur, India
Kanika Soni	National Institute of Technology, Delhi, India
M. Kiruthiga Devi	Department of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India
Meet Oza	Manipal University Jaipur, Jaipur, India
Monika Agarwal	Dayanand Sagar University, Bangalore, Karnataka, India
Manish Dixit	Department of Computer Science and Engineering, Madhav Institute of Technology and Science, Gwalior, Madhya Pradesh, India
Nitesh Pradhan	LNM Institute of Information Technology, Jaipur, India
Neha Kohli	GD Goenka University, Gurugram, India
M.R. Pooja	Vidyavardhaka College of Engineering, Mysuru, Karnataka, India
Ritik	GB Pant DSEU, Okhla-1 Campus, Delhi, India
Rajni	GB Pant DSEU, Okhla-1 Campus, Delhi, India
Sushma Hans	Amity University, Dubai Campus, Dubai, United Arab Emirates
Sunita Kumari	GB Pant DSEU, Okhla-1 Campus, Delhi, India
Shradha Dubey	Department of Computer Science and Engineering, Madhav Institute of Technology and Science, Gwalior, Madhya Pradesh, India
Shivani Batra	KIET Group of Institutions, Delhi-NCR, Ghaziabad, Uttar Pradesh, India
Shelly Sachdeva	National Institute of Technology, Delhi, India
T. Lurthu Pushparaj	Department of Chemistry, Tirunelveli Dakshina Mara Nadar Sangam College, Tirunelveli, Tamil Nadu, India
Vijaypal Singh Dhaka	Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India
Vinay Kumar	KIET Group of Institutions, Delhi-NCR, Ghaziabad, Uttar Pradesh, India
Vaishali Arya	GD Goenka University, Gurugram, India

CHAPTER 1

Role of Federated Learning in Healthcare: A Review

Geeta Rani[3], **Meet Oza**[1], **Heta Patel**[1], **Vijaypal Singh Dhaka**[3,*] and **Sushma Hans**[2]

[1] *Manipal University Jaipur, Jaipur, India*

[2] *Amity University, Dubai Campus, Dubai, United Arab Emirates*

[3] *Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India*

Abstract: In the modern era, there is a boom in automating medical diagnosis by adopting emerging technologies and advanced applications of artificial intelligence. These technologies require a huge amount of data for training the models and precisely predicting the disease or disorder. Multiple organizations can contribute data for such systems but maintaining data privacy while sharing the data is a major challenge. Also, provisioning a large data corpus for the performance improvement of machine learning and deep learning models in the healthcare domain while keeping the patient's medical confidentiality intact is a point of concern. Thus, there is a strong need to preserve the privacy of medical data. This calls for the use of up-to-the-minute technologies where the necessity of sharing raw data is completely eradicated, while each organization receives a catered infrastructure for processing data. A cross-silo federated learning model is based on the concept of decentralized data weights collection from multiple clients which are then processed on the central server for modeling and aggregation, thus maintaining data privacy in its true sense. The authors in this manuscript provide a detailed comparative study of the different deep learning-based models in federated learning and how efficiently they can classify lung X-Ray images into three classes: Covid-19, Pneumonia, and Normal. This study can provide a benchmark for the researchers looking forward to deep learning-based model applications of cross-silo federated learning in healthcare.

Keywords: Covid-19, Diagnosis, Deep learning, Federated, Medical, Machine learning, Segmentation, X-Ray.

* **Corresponding author Vijaypal Singh Dhaka:** Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India; E-mail: vijaypalsingh.dhaka@jaipur.manipal.edu

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

INTRODUCTION

When Covid-19 pandemic hit the world, it became very important to figure out various ways to detect the existence of the novel virus, apart from the usual RT-PCR tests. Scientists and researchers around the globe have studied, developed, and presented numerous ways to detect the novel Covid-19 virus. Many researchers have also provided multiple ways to detect Covid-19 and pneumonia through CT-Scans and X-rays of lungs with the use of machine learning techniques. But the accurate prediction *via* these machine learning and deep learning models for detection requires a large amount of dataset for training the models. In real scenarios, such large datasets are either not feasible due to system constraints or the patient's medical confidentiality gets impaired. The feasibility issues persist even though we have access to many image annotation tools available in the market. This is because such tools and services are very expensive, and they also require expert supervision and proficiency especially if they are utilized for disease diagnosis. Along with monetary hindrances, the feasibility issue also refers to the communication overhead that would occur due to a large dataset being transmitted for centralization [1]. Such an issue exists in the healthcare domain and across all domains where the models are required to learn and train with the help of multiple clients' data without invading the users' privacy.

The application of traditional machine learning and deep learning in the field of healthcare has already been studied by various researchers for tumor prediction [2], covid-19 screening [3, 4], disease prediction [5], cardiovascular diseases prediction [6], coronary artery disease prediction [7], diabetes prediction [8], glaucoma detection [9], *etc.* A similar case pertaining to users' privacy was encountered by Google in 2016 when the team coined the term 'Federated Learning' while advocating an advanced and novel approach that utilizes distributed data from mobile devices for training. Further, this approach presents how a central model is updated by only using the aggregate of the parameters of the local mobile devices [10]. Federated learning inherently trains the central model based only on the parameters passed on by local machine learning models. In addition to only sharing the parameters, the parameters are also encrypted before being passed on which increases data privacy. Federated learning can be differentiated from distributed learning because of the fact that the main objective of federated learning is training on a large dataset from different clients without the transfer of raw data. Whereas distributed learning focuses on distributing the computing resources across clients [11]. Incorporation of this Federated Learning along with cross-silo transferred learning opens new doors to endless possibilities of more accurate innovations due to the availability of a huge data corpus for training without actually having to exchange or transfer the data. In cross-silo federated learning, data is segregated as silos, *i.e.*, multiple confined data sources which in turn centrally aggregate and train a model by passing out only the trained weights and parameters from each client. For collaboration between institutions of healthcare, finance, *etc.* user data is extremely sensitive, and open alliances might expose such sensitive data to various vulnerabilities. Cross-silo federated learning only sanctions weights and parameter transfer and hence the data fenced within the silos itself. And therefore, cross-silo federated learning serves as a superior alternative to the traditional centralized machine learning approaches. Federated learning has been so far applied to various healthcare applications. For example [12], distributed learning has been used to solve a problem related to hospitalizations due to cardiac cases; and [13] leveraged machine learning in a federated setting to predict fatality and duration of stay at the hospital using electronic medical records.

Federated learning systems seem promising but due to their nature of a dispersed framework, it faces certain challenges too such as, communication cost, resource cost, security of communication, *etc.* High communication costs refer to the overhead incurred due to ample transmissions of parameters for the training process. Frequent transfer of parameters is required between silos in order to present potent results and hence this communication overhead acts as a hindrance to federated learning especially when the connection is slow, and a considerable number of devices or organizations are involved. Also, since FL works on distributed systems, each system involved might have different computational power, different storage capabilities, and different bandwidths. A single slightly less efficient system can be a weak link to the entire process and on the other hand, providing all the concerned organizations with full-fledged resources might increase the system cost. Apart from these overhead challenges, privacy concerns also prevail in federated learning. Although federated learning is known as a mechanism to preserve user data privacy, it is not general wisdom that FL by itself doesn’t protect data privacy. Recent studies [14] have revealed that as models communicate constantly for the transfer of parameters, the process is seen to be leaking some information in the course. For example, [15] a study showed even a small section of original gradients may be enough to let local data slip from the system. Moreover, since the parameters are obtained *via* model training at the local level, vulnerabilities such as model inversion or attacks on model parameters can corrupt aggregate inference.

Even after weighing the opportunities and hindrances presented by the federated learning approach in healthcare, we can conclude that federated learning still races way ahead of traditional machine learning practices. The user data confidentiality issues that come along with the traditional approaches directly make the patients’

CHAPTER 2

Role of Artificial Intelligence in 3-D Bone Image Reconstruction: A Review

Nitesh Pradhan[3], **Vijaypal Singh Dhaka**[1], **Geeta Rani**[1,*] and **Monika Agarwal**[2]

[1] *Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India*

[2] *Dayanand Sagar University, Bangalore, Karnataka, India*

[3] *LNM Institute of Information Technology, Jaipur, India*

Abstract: Three-dimensional geometry of a bone is important in the correct diagnosis of a disease, arthritis, or other bone deformities. The modalities such as Computer Tomography Scans and Magnetic Resonance Imaging are used for a three-dimensional view of a bone. Both the above- stated modalities have high costs and expose the patient to strong carcinogenic radiations. Computer Tomography captures an extensive number of images to collect the required information from a bone. Another modality Magnetic Resonance Imaging is more suitable for retrieving information from soft tissues rather than bones. Therefore, it becomes less effective to read the pathology from bones. This has motivated the authors to identify imaging techniques useful in detecting the pathology or deformity in bones. Also, this is the need of the hour to provide a low cost and safer technique of bone imaging. To address this need, we present a review of the bone imaging techniques and techniques applied for the conversion of two dimensional images into three-dimensional form. We also give the directions for developing the patient-specific and organ-specific optimized techniques for 3-D reconstruction.

Keywords: Dual-energy, Deep learning, Deformation, Femur, Machine learning, Three dimensional, X-ray.

INTRODUCTION

Three Dimensional (3-D) structure of an organ is important for the diagnosis of a disease. The 3-D structure provides an exact configuration and orientation of the organ. So, it is useful in the detection of disease in tissues or fractures in bones.

* **Corresponding author Geeta Rani:** Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India; E-mail: geetachhikara@gmail.com

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

This structure also gives information about chronic bone loss such as a glenoid defect in recurrent shoulder dislocation and the extent of osteophytes in the arthritic joint. Computer Tomography (CT) scan and Magnetic Resonance Imaging (MRI) provide a 3-D configuration of a bone. But these techniques are less preferred than 2-Dimensional (2-D) techniques due to their high cost and heavy exposure to carcinogenic radiations [1].

The Dual-Energy X-ray Absorptiometry (DXA) images are used to diagnose osteoporosis in patients. In the DXA image, the value of the T-score is presented. The value of the T-score between +1 to -1 indicates healthy bone. Its value between -1 to -2.5 shows that the bone has become prone to osteoporosis [2]. This state is named osteopenia. A value below 2.5 is an indication of the poor quality of a bone and a sign of osteoporosis. In this disease, there is a decrease in Bone Mineral Density (BMD). Thus, it increases the risk of bone fracture. As per details given in [3], in Europe, 30% of women of the age of 50 years suffer from osteoporosis. As per the report published in the year 2000, 3.1 to 3.7 million cases of osteoporosis were recorded. The direct cost paid for the treatment of this disease was reported 32 billion [4]. As per the trend reported in a study [4], the cost may rise to 76.8 billion per year in 2050.

X-ray imaging is a medical imaging technique used to capture bone deformities. This technique shows a clear demarcation between the bones and soft tissues. So, it becomes easy to read the information about bone deformity if any. X-ray imaging is preferred over a CT scan in case of weight-bearing imaging and dynamic imaging of joint motion (fluoroscopy). People prefer X-ray imaging due to its high availability and low cost. Therefore, there is a requirement of proposing a technique that can provide a 3-D view of a bone using 2-D imaging techniques such as DXA and X-rays. The differences among the above-stated medical imaging techniques are given in Table **1**. The techniques for showing 3-D images of bones or ways to reconstruct 3-D images from 2-D images are shown in Fig. (**1**).

Table 1. Difference between CT scan, X-ray and DXA techniques.

Parameters	CT Scan	X-ray	DXA
Radiation Vulnerability	The effective radiation dose from CT ranges from 2 to 10 millisievert (mSv) [5].	Exposure to ionizing radiation is about 0.1mSv [5].	Exposure to ionizing radiation lies in the range of 0.1μSv to 5μSv [6].
Cost	Cost of CT Scan lies in the range from $1,200 to $3,200 [7].	The cost is from $1200 to $4000 [7].	This is three times more costly than the X-ray technique [7].

(Table 1) cont.....

Parameters	CT Scan	X-ray	DXA
Time taken for a complete scan	Takes about 5 minutes.	Takes a few seconds.	Takes more time than X-ray techniques but less time-consuming than the CT scan technique.
Application	Suitable for bone injuries, Lung and Chest imaging, and cancer detection.	Useful to examine fractured bones and to diagnose diseases in tissues.	Useful to diagnose osteoporosis or measure mineral density of bone.

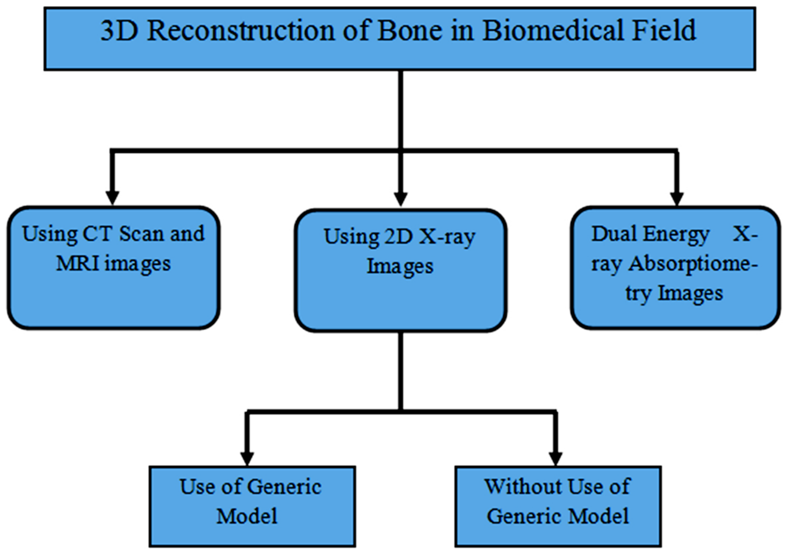

Fig. (1). Imaging in biomedical field images.

In this chapter, the authors present a review of 3-D and 2-D techniques used in medical imaging as discussed in Table **1**. This chapter also gives a comparative analysis of techniques used for the 3-D reconstruction of bones from 2-D imaging techniques.

ANALYSIS OF RELATED WORK

A review of related works shows that the 3-D reconstruction of medical images from 2-D images attracted researchers to propose different models. The contributions from the researchers are discussed below.

WEI *et al.* [1] identified that X-ray imaging is preferred over the CT scan image due to the low intensity of exposure and low cost. Therefore, they proposed the 3-D recreation system for the femoral shaft shape. They used an orthographical heading for recreating a 3-D femur. They used a numerical morphology strategy

CHAPTER 3

Role of Machine Learning and Deep Learning Techniques in Detection of Disease Severity: A Survey

Geeta Rani[1], **Vijaypal Singh Dhaka**[1*] and **Sushma Hans**[2]

[1] *Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India*

[2] *Amity University, Dubai Campus, Dubai, United Arab Emirates*

Abstract: The increasing number of health issues is a cause of concern for public as well as health services across the globe. However, a boom in the use of imaging techniques such as CT scans and chest radiographs has been observed for correct diagnosis. But, manual scanning of these modalities requires expertise in modality reading. It is also a time-consuming task. Artificial intelligence-based techniques have proven their potential in pattern recognition, object identification, and data analysis. Therefore, these techniques can be used to provide assisting tools for the primary screening of diseases from these modalities. It has been observed from the literature that a lot of research works are available on disease diagnosis and classification using machine learning, and deep learning. But, the disease severity detection is underexplored. Moreover, the techniques employed for the detection of the severity of diseases have lacunae that need immediate attention. These challenges motivated us to review the machine learning and deep learning-based technological solutions proposed in the literature for the detection of disease severity. The objective of this research is to present a comprehensive survey of research works available about disease severity detection. This research also presents a comparative analysis of the machine learning techniques and deep learning techniques employed, datasets used, and performance achieved. It also highlights the drawbacks of the technological solution proposed. Further, it provides the directions for future scope in the domain of disease severity detection.

Keywords: Artificial intelligence, Disease, Deep learning, Machine learning, Severity.

[*] **Corresponding author Vijaypal Singh Dhaka:** Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India; E-mail: vijaypalsingh.dhaka@jaipur.manipal.edu

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

INTRODUCTION

The onset of COVID-19 in the year 2020 enforced us to rethink about the importance of individual as well as community's health. Also, it worked as a driving force to investigate the quick and effective means of disease diagnosis. Thus, a huge change in the health infrastructure, disease diagnosis techniques, and human resource management is observed. The technological solutions have been proposed as an alternative to traditional laboratory tests to minimize the cost of disease diagnosis and delay in report generation.

The study of the literature reveals the potential of Machine Learning (ML) and Deep Learning (DL) techniques in image quality enhancement [1, 2], pattern recognition, and image reconstruction [3, 4]. Thus, these techniques have been widely used for disease screening in human beings [5 - 7] as well as plants [8].

In recent years, the detection of the severity of diseases has gained significant attention from the research community. Detection of the severity of a disease at an early stage is essential to reduce its impact on the patients and mortality risks in patients. Many patients may miss the correct time of treatment as only a few symptoms appear at an early stage of disease. Therefore, there is a need to utilize the potential of computer-aided automatic diagnostic methods for detecting disease severity.

The degree of accuracy of severity prediction using traditional laboratory tests is highly dependent on the methods of sample collection, sample storage, transportation, and equipment used. Also, these techniques are inconvenient to patients, incur high costs, and lead to delays in severity detection. Thus, it may increase the mortality rate. These challenges motivated us to explore the technological solutions available for the detection of disease severity.

The comprehensive study of available research works shows a significant contribution from the research community in the arena of machine learning and deep learning for disease severity detection. Indeed, various ML techniques have been employed for severity detection when parametric data is available in the form of reports, ECG signals, clinical features, *etc.* [10 - 16]. The researchers employed the techniques such as K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Random Forest (RF), Multi-tree XGBoost model and AdaBoost. The comparative analysis of these techniques is shown in the subsequent section in Table **1**. These techniques have their own significance based on the type of dataset available and the result required. However, ML techniques reported the highest accuracy of severity detection approximately 100% but there is a need to improve the reliability and robustness of these techniques. Moreover, these techniques require clinical data or data collected in terms of reports. The data collection mechanisms are inconvenient to patients, time-consuming, and require high cost. Thus, there is a need to integrate sensor-based data collection, transmission to cloud storage, and implementation of ML models on the data stored on cloud storage for disease diagnosis and severity detection.

Table 1. Comparative analysis of existing DL-based techniques used for disease detection and severity prediction.

Authors, Year, and Citation	Technique Applied	Dataset used	Performance	Challenges
Yao *et al*., 2020	Predictive Logistic Regression Support Vector Machine, Random Forest, K-Nearest Neighbour (KNN), and AdaBoost	Health parameters recorded by urine and blood test of patients of COVID-19, and their clinical features such as age, blood pressure, heart rate *etc.*	Severity prediction accuracy of 99.17% on the training dataset, and 81.48% on the testing dataset.	Conclusions made using a small dataset confined to only one geographical location. Collection of blood, and urine sample is required. High cost is involved in blood and urine tests. Time consuming, Accuracy is dependent on sample collection, storage, and equipment's available in laboratories.
Li Yan *et al*., 2020	Multi-tree XGBoost model for categorizing critical and severe cases of COVID-19	Selected three features viz. Lactic Dehydrogenase Lymphocyte, and high sensitivity C-reaction protein from 375 patients of COVID-19	Predicted the survival rate with the accuracy of approximately 90%	Complete dataset collected from only one geographical location. Prediction made only on the basis of three parameters. High cost and time-consuming due to laboratory tests.
Tadesse *et al.,* 2020	SVM	ECG signals from patients of Hand Foot and Mouth Disease(HFMD) and tetanus.	98.1% accuracy for HFMD, and 78% for the tetnus.	The quality of ECG is dependent on the movement of the patient during monitoring. Tested on a small dataset of only 10 patients.
Tripoliti *et al.,* 2017	Review of KNN, SVM, for prediction of severity of heart disease and heart rate failure	PhysioNet dataset	Best accuracy of 96.39% by using KNN	-
Liu *et al.*	KNN, and SVM,	PhysioNet dataset	Claimed accuracy of 100% for KNN	Problem of class imbalance in the dataset, trained and tested on very small dataset.

CHAPTER 4

Computer-aided Bio-medical Tools for Disease Identification

E. Francy Irudaya Rani[1], **T. Lurthu Pushparaj**[2] and **E. Fantin Irudaya Raj**[3,*]

[1] *Department of Electronics and Communication Engineering, Francis Xavier Engineering College, Tirunelveli, Tamil Nadu, India*

[2] *Department of Chemistry, Tirunelveli Dakshina Mara Nadar Sangam College, Tirunelveli, Tamil Nadu, India*

[3] *Department of Electrical and Electronics Engineering, Dr. Sivanthi Aditanar College of Engineering, Thoothukudi, Tamil Nadu, India*

Abstract: The health expert's crucial task is to interpret the output and treat the disease accordingly. They may delay the decision-making during emergencies. To address this issue, research on smart tools for biomedical applications is much needed which may help in making accurate decisions at the earliest stage. Discovery in medicinal research requires state-of-the-art computer-based tools for diagnosing and treating complex diseases such as cancer, COVID-19, SARS-Cov, MERS-Cov, tuberculosis, brain disorders, heart, and lung-related chronic infections. Among various diagnostic methods, image-based disease identification stands out as the most prominent approach for detecting new and complex diseases. A well-trained computerized biomedical system can provide physicians with enhanced support for early disease detection. Biomedical images are typically acquired from various sources, including CT, ultrasound, MRI, dermoscopy, X-ray, biopsy, and endoscopy. Presently, a wide range of image-analysis procedures are available for biomedical images. These procedures involve image acquisition, pre-processing, segmentation, feature extraction, and classification, all contributing to improved disease decision accuracy. Although many biomedical images are available online free of cost, the proper procedure must be followed to select appropriate images from databases and enhance their quality. This is important for effectively training image-processing algorithms and increasing their efficiency. This leads to improved instrument performance and more valuable insights into the diseases under study. It also handles complex and vast image data to detect early signs of unusual signals, growth, inflammation, cell damage, protein sequence changes, and blockages. Additionally, it should be user-friendly and convincing to health experts to identify hidden biological issues. This chapter emphasizes the power of computerized tools in image analysis and disease detection. It also focuses on recent developments in the field of medicinal research.

[*] **Corresponding author E. Fantin Irudaya Raj:** Department of Electrical and Electronics Engineering, Dr. Sivanthi Aditanar College of Engineering, Thoothukudi, Tamil Nadu, India;
E-mail: fantinraj@gmail.com

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

Keywords: Biomedical image processing, CAD, Combined algorithm, Cancer diagnosis, Covid-19, Classification, Feature extraction, Pre-processing, Segmentation.

INTRODUCTION

Computer-based medical image analysis is a crucial task for physicians in diagnosing new diseases. In biomedical research, computer-assisted detection (CAD) has become widespread, even generating its own field of study. However, it cannot replace a physician, it serves as an assisting device in making accurate and quick decisions. It may prove an asset especially during medical emergencies. CAD can be broadly categorized into CA detection and CA-Diagnosis. CA detection involves marking visible parts in images, while CA-Diagnosis is used to analyze the identified structures.

Deep Learning (DL) exhibits the ability to perform pattern recognition, image processing, pattern learning, object detection, and pattern matching. Thus, it is a valuable tool for early detection of tissue damage and human genome disorders [1]. CAD is not a single platform; rather, it is a combination of intelligent techniques like machine learning, machine vision, and medical imaging, forming a hybrid model. It encompasses various techniques, including PET, MRI, Dermoscopy, Biopsy, X-rays, mammography, ultrasound, and CT Imaging. Till now, nearly one million images have been captured and analyzed.

CAD gained popularity in the analysis of stroke caused by brain injury and abnormalities in abdominal images. It has the potential to translate blurry or unclear medical images into high-contrast ones. Additionally, it can reconstruct and enhance images or the region of interest (ROI). Moreover, CAD enables the generation and tracking of medical reports for patients. These applications have facilitated the detection of normal and chemically altered COVID-19 infections in patients. Primarily, these procedures are utilized for the early detection of tumors with the potential to develop cancerous tissues.

The remarkable output accuracy from AI, DL, and ML methods helps doctors and patients by reducing the risk of exposure to hazardous ionizing radiation from two-dimensional (2-D) and three-dimensional (3-D) imaging instruments [2].

APPLICATIONS OF CAD IN MEDICAL ANALYSIS

CAD approaches are helpful in recognising various diseases ranging from fever to life-threatening diseases such as AIDS, cancer, Covid-19 (Fig. **1**) outbreaks, heart attack, brain injury, skin abnormalities, and diabetic retinopathy. This tool is effective in assessing chronic health conditions before the onset of visual symptoms.

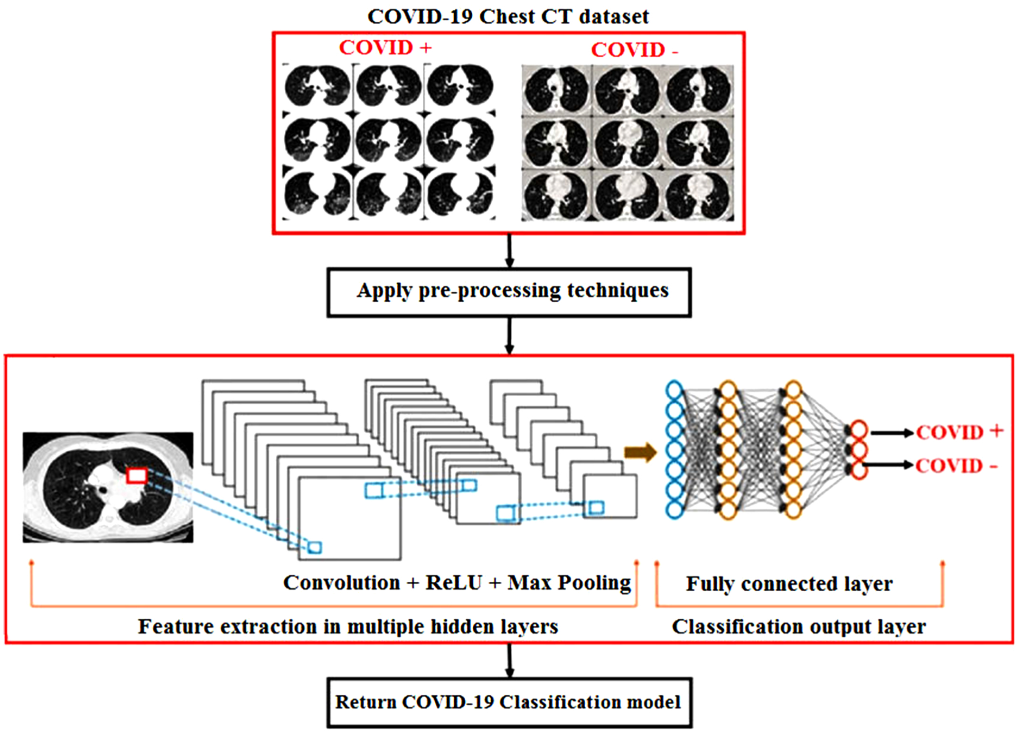

Fig. (1). Application of CAD tools in COVID-19 infection classification.

Cardiology Study using CAD

Echocardiography is the most used screening tool in cardiovascular medicine. It has a simple data-capturing and interpretation mechanism. Thus, it can be integrated with CAD to provide a cost-effective, radiation-free procedure. Furthermore, 3-D imaging techniques such as CT and MRI are employed to give a comprehensive heart anatomy.

CAD as shown in Fig. (**2**) can preprocess the echocardiography dataset and provide superior output for echocardiographers. The field of CAD has recently made significant progress and introduced the Echo Net, a deep neural network. The model is capable of distinguishing cardiac structures and assessing their activity. Echo Net also predicts systemic symptoms that may not be easily recognized through probabilistic reasoning. Additionally, information echocar-

CHAPTER 5

Prognosis of Dementia using Machine Learning

Anu Saini[1], **Sunita Kumari**[1,*], **Ritik** [1], **Rajni** [1] and **Sushma Hans**[2]

[1] *GB Pant DSEU, Okhla-1 Campus, Delhi, India*

[2] *Amity University, Dubai Campus, Dubai, United Arab Emirates*

Abstract: The brain is one of the most sensitive parts of the human body which transmits millions of signals every moment. Dementia is the most emerging brain health issue which involves memory loss, difficulty in problem-solving, handling complex tasks, *etc.* Dementia is a syndrome that causes a loss of mental ability. It affects memory, thinking, shape, comprehension, counting, reading ability, language, and judgment. Dementia affects millions of people and can be the leading cause of death. It is now the seventh leading cause of death worldwide, as well as one of the major causes of impairment and reliance on elderly people. There is no treatment for dementia at present. The importance of early detection and diagnosis in improving early and effective management is crucial. Predicting dementia in advance can lead us to a better life. To predict dementia, various Machine Learning models have been used. In this paper, Dementia is predicted on the basis of MRI Images, for this, three different datasets of MRI Images have been collected. Furthermore, for better prediction, various Machine learning models are used to predict dementia and validate the performance with statistical analysis like K-Nearest Neighbours, XG Boost, Support Vector Machine, Random Forest Algorithm (RFA), and Convolutional Neural Network (CNN). Out of all algorithms, Random Forest Algorithm and Convolutional Neural Network gave the best result with the accuracy of 93.2 and 99.9 respectively.

Keywords: Alzheimer's disease, CNN (Convolutional Neural Network), Dementia, Random forest algorithm.

INTRODUCTION

Dementia is a neuropsychological disorder that causes memory loss leading to paralysis and reliance on others for survival. It is most frequent in those over 60 years, with an estimated 900 million people in their 60s suffering from it. It affects more than 46 million individuals globally, more than the whole populace of Spain. By 2050, this number is predicted to surge to 131.5 million [1]. Due to their severe mental incapacity, Alzheimer's disease and other major dementias

* **Corresponding author Sunita Kumari:** GB Pant DSEU, Okhla-1 Campus, Delhi, India;
E-mail: sunita2009@gmail.com

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

affecting older people inflict a substantial financial and social burden on families and communities, with Alzheimer's account for over 75% of cases.

Dementia affects approximately 50 million people worldwide, as per the World Health Organization (WHO), with an estimated 10 million new cases expected annually.

Dementia is clinically diagnosed based on a comprehensive health history provided by patients and families, as well as neurological tests followed by a cognitive assessment. Tests such as hematology, CT, and MRI should be performed to diagnose the cause of dementia. Neuropsychological exams are important for identifying inefficiencies in human “cognitive domains”. Although there are a few clinical steps for the early detection of dementia, there are still many challenges.

Once it is diagnosed, currently there is no viable medication for dementia that can halt or stop its growth. Hence, it becomes critical to concentrate on the early phases, prompt intervention, and disease prevention. Timely diagnosis can decide the degree of dementia, and neuro-imaging analysis can aid in your diagnosis.

Further, it is better if we can predict whether a person can be affected by dementia in his/her future years. The condition is predicted using a variety of machine-learning technologies. This research aids in determining how Deep Learning can be used to predict dementia using MRI pictures.

We employed two methods to achieve the highest accuracy of our model, Convolutional Neural Network (CNN) for image-based analysis and RFA-based statistical analysis.

CNN is a robust image processing and artificial intelligence (AI) system that uses in-depth learning to execute productive and descriptive functions, often combining image and video recognition and complimentary algorithms as well as natural language understanding (NLU). The neural network is a hardware and/or software system that controls the pattern in the human brain once neurons are activated. Traditional neural networks aren't well adapted to image processing, thus they should be provided with fragmented images. CNN has its own incredibly sophisticated “sensors” similar to those used in the old record, a site dedicated to human and animal visual processing. To solve the difficulty of standard neural networks processing images, the layers of neurons are structured in such a way that they span the full observing field.

CNN employs a multilayer perceptron technology with low processing requirements. CNN layers comprise several flexibility layers, integration layers, completely integrated layers, and custom layers, as well as an input layer, an output layer, and a hidden layer. The removal of image processing limitations and greater efficiency results in a more efficient, simple-to-use training system that limits image and natural language processing.

Random Forest is a machine-learning strategy for dealing with setbacks and planning issues. It employs integrated learning, which is a multidisciplinary method to solve complex issues. The random forest algorithm comprises a large number of trees that can be pruned. Combining bags or bootstraps is used to train a 'forest' formed by a random forest algorithm. Bagging is a set of meta-algorithms that help machine learning algorithms increase their accuracy. The result is determined by the algorithm (random forest), which is based on decision tree prediction. It makes predictions based on the rate of output of different trees. The accuracy of the output improves as the number of trees grows.

The random forest overcomes the decision tree's algorithmic limitations. It reduces data set overload and improves accuracy. Predictions are generated without the requirement for various package configurations (like sci-kit-learn).

Random Forest Algorithm Characteristics:

• Provides an effective technique to handle missing data.

• More accurate results as compared to the decision tree algorithm.

• Can provide sound guesses without adjusting the upper parameter.

• Resolves decision tree overcrowding.

The rest of the paper is structured as follows:

The related work done by various researchers is found in section two. Section three discusses the methodology proposed by the authors. Section four describes the experimental work and result analysis. Finally, in section five, there is a conclusion.

RELATED WORK

Literature shows that lots of researchers used different machine learning models to predict dementia. Some of the works are discussed in detail below.

The authors of the paper [1] gathered data from 5,272 people who completed a survey having 37 items. They focused on the three alternative ways of selecting features in order to find the most significant ones. The diagnostic models are then

CHAPTER 6

A Clinical Decision Support System for Effective Identification of the Onset of Asthma Disease

M.R. Pooja[1,*]

[1] *Vidyavardhaka College of Engineering, Mysuru, Karnataka, India*

Abstract: We present a clinical decision support system for the identification of asthmatics in two different cohorts representing rural and urban populations in India. The input data representing the two populations are cross-sectional in nature and are necessarily categorical in nature, with information on clinical history emphasizing clinical symptoms and patterns characterizing the disease. The system is described as hybrid as it combines the unsupervised and supervised learning techniques in a unique way as discussed in the work presented in the paper. The clustering information emphasizing the phenotypic characterization of asthma is an input to the classifier and a significant improvement is observed in the performance of the classifier. The results of the developed hybrid decision support system are quite promising for suitable deployment in a real-time scenario, as it explores the benefits of both supervised and unsupervised learning techniques. Further, the use of clustering information in the form of cluster evaluation scores as an input parameter to the classifiers can efficiently predict disease outcomes, especially with diseases such as asthma, as the disease is heterogeneous and exhibits several disease subtypes and heterogeneous phenotypes.

Keywords: Correlation, Hybrid, ISAAC, MFCM, Subject clustering.

INTRODUCTION

Asthma, being one of the most chronic respiratory diseases of the lungs is known to be affected by various factors including, genetic and environmental features that play a vital role in the persistence and progress of the disease [1]. Clinical symptoms and their patterns along with their frequency have a great impact on the disease outcome and impact when assessed using electronic health record data [2]. ISAAC is a well-recognized body that has formulated core questionnaires with respect to asthma and its comorbidities that have been received well in most of the pilot studies conducted across several regions. A hybrid decision support system that operates in two stages has been proposed, wherein the results generated dur-

[*] **Corresponding author M.R. Pooja:** Vidyavardhaka College of Engineering, Mysuru, Karnataka, India; E-mail: pooja.mr@vvce.ac.in

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

ing the first stage are effectively used as a significant input in the subsequent age [3]. A significant fact contributing to the improvement in the prediction of class outcomes is the use of clustering information during the process of classification [4]. This is seen as a requisite to the system as asthma is seen to exhibit polymorphic phenotypes that have their own characterization [5].

Identification of patient populations at high risk is an important intervention in the early detection and clinical assessment of chronic diseases like asthma and bronchitis, as it can lead to targeted and personalized therapies. Our research has a credible potential to be an important intervention through facilitating early risk assessment and effective prognosis of the disease *via* risk stratification. In case of failure to recognize the disease at an early stage, the disease could be progressive and finally become irreversible, though reversible in nature. This necessitates us to propose approaches that help us in the prediction of the disease using severity indicators that can be possibly identified at an early stage, enabling early medical interventions that alleviate the disease severity, subsequently lowering the mortality rate. The disease is further heterogeneous, in that it presents itself with variable overlapping syndromes of other obstructive diseases, which makes it difficult to delineate the symptoms characterizing asthma.

RELATED WORK

Clinical Decision Support systems incorporating machine learning techniques to mine the useful data may be adopted to explore asthma data, and to identify latent patterns that add value, while providing a supplementary source of better understanding for decision-making with respect to risk factor identification along with the assessment of the severity of the disease and the extent to which it relates to other diseases of the lungs such as COPD (Chronic Obstructive Pulmonary Diseases) [6 - 8].

Bayesian classification approaches implementing supervised learning techniques have resulted in classes that are principally identified as correct and have a good amount of face validity along with substantial relationships to asthma, variations in airway reactivity and lung function, which illustrate content validity with a good amount of gratification [9, 10]. Predictive analytics have been suitably deployed to estimate risk factors by exploring a blend of independent variables from variable sources [11]. Rule-based systems integrating expert-driven feature selection and rough sets techniques have been identified to describe asthmatic populations visiting emergency departments [12]. A good amount of precision was obtained with the task associated with the differentiation of asthma against other respiratory disorders depending on the nature of the underlying data and the ensemble learning techniques used [13].

MATERIAL AND METHODS

Dataset Description

The response data for the ISAAC core questionnaires gathered under the first phase, a study that has been used across several pilot studies, was used to validate the system performance [14]. The questions used in the collection of data typically signify sensitive and specific indicators of the disease. The data are cross-sectional in that it involves the analysis of data concerning a specific population which is widely a significant subset that is utmost representative at a specific point in time. The data was gathered across two age groups, 13 to 14 years and 6 to 7 years. While in the older group of children, the response was completed by the children themselves, and the same was completed by the parents in the other group.

Combatting Class Imbalance

The New Delhi dataset contained the details of 2961 subjects explained with respect to 47 variables indicative of asthma comorbidities and symptoms. The disease outcome was reported as unknown for 112 samples. Hence, we excluded all 112 samples yielding a dataset containing 2849 samples. However, only 110 subjects out of the total population were asthmatics and hence we sought to combat the imbalance by performing stratified sampling that was able to draw 110 samples from the other category too (non-asthmatics), thus ending up with a balanced dataset that was fed as the input data. Along similar lines, we worked to draw 2% amounting to 64 subjects from the non-asthmatics from the total population to combat the class imbalance that would have resulted in 78 asthmatics in the population, resulting in an input dataset containing 142 subjects.

Feature Clustering

Initially, the predominant features that can be identified as significant comorbidities of asthma disease are chosen by applying feature clustering. Feature clustering is performed by employing Modified Fuzzy C means clustering (MFCM), which uses an objective function based on correlation, unlike the traditional Fuzzy C Means clustering which uses an objective function based on distances [10, 11]. We chose to organize the input feature set into four clusters as per the evaluation results of subtractive clustering. The features that coexist with asthma attributes are extracted from the cluster containing asthma. This leads to a reduced feature set in the process.

A correlation-based objective function is used in the fuzzy c means clustering. The distance metric here incorporates a Pearson's correlation coefficient that is

CHAPTER 7

Applying Deep Learning and Computer Vision for Early Diagnosis of Eye Diseases

Shradha Dubey[1,*] and **Manish Dixit**[1]

[1] *Department of Computer Science and Engineering, Madhav Institute of Technology and Science, Gwalior, Madhya Pradesh, India*

Abstract: Medical image processing has a significant role in clinical investigation and recent medical research. An appropriate image-based medical assessment helps to analyze or detect critical diseases early, as it has a high value of medical information. In this study, medical imaging is reviewed for the diagnosis of eye diseases using computational intelligence. However, the identification of these diseases using traditional image processing is quite complicated. Nowadays, various machine learning and deep learning approaches are developed for the detection of different eye diseases which are helpful for the detection of the diseases at an early stage. Research showed that eye disorders are more serious in emerging or underdeveloped nations due to inadequate healthcare facilities and skilled health workers. An estimate of 45 million people around the world are blind and the tragic fact is that only 75% of these cases are curable. Moreover, the doctor-patient ratio around the globe is about 1: 10,000. Therefore, it takes an hour to create a screening system for the identification of these illnesses. Ophthalmology is close to making breakthroughs in evaluating, diagnosing, and treating eye diseases. Additionally, many eye and vision problems show no obvious signs. As a consequence, people are often unaware that problems exist. Early detection of diseases is a primary concern as they could be easily cured before leading to severity. This research paper focuses on detecting eye illnesses, such as Diabetic retinopathy, Diabetic Macular Edema, Glaucoma, Age macular Degeneration, Retinal Vascular Occlusions, and Retinal Detachment. The authors explore various algorithms, imaging modalities, and challenges in this context. The study aims to raise awareness about eye disorders leading to blindness using computer vision, image processing, and deep learning techniques. It also investigates how these machine learning and deep learning approaches can aid in early disease diagnoses for effective treatment before vision loss occurs.

Keywords: Age-related macular degeneration, Cataract, Deep learning, Diabetic retinopathy, Glaucoma, Heidelberg retinal tomography, Optical coherence tomography, Ultrasound imaging.

* **Corresponding author Shradha Dubey:** Department of Computer Science and Engineering, Madhav Institute of Technology and Science, Gwalior, Madhya Pradesh, India; E-mail: dubeyshradha29@gmail.com

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

INTRODUCTION

Clinical research and contemporary medical research both benefit from medical image processing. An appropriate image-based medical assessment helps to analyse or detect critical diseases early, as it has a high value of medical information. This study presents a review of medical imaging for diagnosing eye diseases through computational intelligence. Traditional image processing methods have proven challenging for disease identification. However, the introduction of diverse machine learning and deep learning approaches has revolutionized the detection of various eye diseases, enabling timely identification and diagnosis. Further sections in this chapter delve into the details of these approaches and their significant contributions to early detection of eye disorders.

MOTIVATION

Vision is a crucial human sense, but an increasing number of patients are affected by retinal diseases daily. A healthy retina is essential for central vision [1]. Detecting these diseases early can protect the retina from damage, but it presents a significant challenge to medical science. With the growing prevalence of ocular diseases due to demographic aging and poor dietary habits, there is a pressing need for effective detection methods. This research paper reviews various computational intelligence techniques utilized to classify healthy and diseased retinal images and categorize the severity of retinal diseases based on ocular image abnormalities. Precise classification of retinal conditions can reduce unnecessary hospital visits, particularly amid the rising cases of infectious diseases. The proposed approach aids in diagnosing the type of retinal disease and assessing its severity level, facilitating timely treatment and improved patient care.

This chapter's primary contribution lies in discussing the effectiveness of deep learning and medical imaging for timely disease detection. Additionally, it explores various methods and techniques to capture eye images using different machines and approaches. This chapter extensively explores various imaging systems used in diagnosing different diseases, each with its advantages and limitations. It also provides a detailed analysis of numerous eye diseases, including their causes, risk factors, and symptoms, aiming to facilitate early diagnosis and proper treatment. The literature discusses several deep neural networks tailored for specific disease detection, enabling swift and accurate diagnoses even in urban or rural settings without the need for specialized professionals. Following detection, patients can promptly seek medical advice for appropriate treatment. These deep learning models offer rapid reports, unlike other cutting-edge methods that may take 2-3 days for results, which could be potentially harmful.

The chapter is structured into several sections. Section 2 presents an introduction to deep learning. In Section 3, a concise review of the existing literature is provided. Section 4 delves into the discussion of various imaging equipment necessary for capturing eye images to compile the dataset for disease diagnosis. Section 5 offers a comprehensive examination of different eye diseases and the applications of computer vision and deep learning in their detection. Moreover, the chapter addresses research challenges in Section 6 before concluding with a summary of key findings.

TECHNICAL ASPECTS OF DEEP LEARNING

The term "computer-aided diagnosis" refers to analyzing relevant clinical data using computer-based filters or tools to identify disease-related patterns [2]. In this study, various deep learning algorithms are explored, particularly in the field of computer vision applied to eye images. For image processing applications [3], Convolutional Neural Network (CNN) models are extensively used and studied by researchers. These CNN models fall into different categories, including LeNet, AlexNet, VGGNet, GoogleNet, InceptionV3, ResNet, and more [4].

Convolutional Neural Networks (CNNs) are deep neural networks used to enhance traditional image processing tasks. They are widely popular due to their reduced pre-processing requirements compared to other algorithms. CNNs effectively capture temporal and spatial dependencies in an image using filters. A typical CNN consists of three main layers: an input layer, hidden layers, and an output layer. The network can have multiple hidden layers, each performing different functions like feature extraction, classification, and feature flattening. The basic block diagram of a CNN is shown in Fig. (**1**), illustrating its application in disease classification. Initially, the convolution function is applied to the input images, where a filter is used to activate specific features. Typically, a non-linear ReLU (Rectified Linear Unit) activation function is employed to make the network robust with various types of input images. Next, the pooling function is used to reduce computation overhead by considering only relevant parameters. Finally, the flatten function converts the spatial dimension into a channel dimension, and a fully connected layer with an activation function, often softmax, is used to obtain the classified image as output.

To construct a robust Deep Learning (DL) model, two key components are essential: the 'brain' (CNN) and the 'dictionary' (the datasets). Initially, CNNs faced computational challenges, but with the advent of GPUs, these methods

CHAPTER 8

The Fusion of Human-Computer Interaction and Artificial Intelligence Leads to the Emergence of Brain Computer Interaction

M. Kiruthiga Devi[1,*]

[1] *Department of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India*

Abstract: A personal computer may be used, with input devices such as a keyboard, mouse, and joystick serving as an interface between the computers and the human. The euphemistic, physically challenged are unable to use these computer systems, therefore, BCI technology has advanced external applications to be managed without physical movements in order to assist these physically disabled people and address the limitations of HCI. The technological advancement in the field of cognitive neuroscience and brain imaging has enabled it to communicate directly with the human brain instead of using an interface. Rather than generating signals from muscle movements, these systems use brain activity to monitor computers or communication devices. Researchers in the field of Human-Computer Interaction (HCI) look at ways for machines to utilize as many sensory sources as possible. Furthermore, researchers have begun to consider implicit types of data, input that is not specifically performed to instruct a machine to perform a task. Systems can evolve dynamically based on this data in order to assist the user with the task at hand. Here we discussed components of Brain-Computer Interface, its characteristics and challenges. The researchers are attempting to replace conventional classifiers with Convolutional neural networks (CNNs) that would provide a promising advantage in classification. The EEG signals from the brain can be linked seamlessly to mechanical systems *via* BCI applications, making it a rapidly growing technology that has applications in fields such as Artificial Intelligence and Computational Intelligence.

Keywords: Actions, Atmosphere, Artificial intelligence, Brain computer interface, Convolutional neural networks, Classifiers, Communication, Computational, Deep Learning, Electroencephalographic, EEG, Human-computer interaction, Input devices, Movements, Machines, Machine learning, Mechanical system, Neuroscience, Physiology, Signals, Sensory, Support vector.

[*] **Corresponding author M. Kiruthiga Devi:** Department of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India; E-mail: kiruthigaprofessor89@gmail.com

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

INTRODUCTION

Sensors are used to track some of the physical processes in the brain associated with specific types of thought [1]. As a result of these developments, brain-computer interfaces (BCIs) have been developed, allowing communication among systems that do not depend on the brain's usual peripheral nerve and muscle systems. Many factors have aided this advancement, including improved understanding of neurobiological processes, machine learning algorithms, and deep learning. A personal computer may be used as an interface between computers and humans with input devices such as a keyboard, mouse, *etc.* [2]. The disabled persons are unable to use these systems, but in order to assist these euphemistic physically challenged people and address the limitations of HCI, BCI technology has advanced, allowing external applications to be managed without the use of physical muscle movements. BCI calibration is difficult with two stages: the low signal-to-noise ratio (SNR) and high subject-to-subject variability. The amount of time needed for calibration varies depending on the type of paradigm used. It can still be reduced by half. If brain signals have a low Signal-to-Noise Ratio (SNR), BCI is most challenging when it comes to accurately determine human intentions [3]. The reality is that BCI's real-world implementation is limited by poor generalization potential and low classification accuracy.

Deep learning approaches have been used to work with brain knowledge in recent years to address the aforementioned challenges. While normally machine learning algorithms need to select features manually, deep learning can learn complex high-level features directly from brain signals, and its accuracy scales well with the size of the training data set [4]. It is possible to monitor the activity of the brain using different techniques, which can be categorized into two types: invasive and non-invasive. The non-invasive BCI system majorly uses Electroencephalogram (EEG) signals to capture brain signals from electrodes mounted on the scalp. EEG data is extremely noisy, therefore, it can be difficult to extract a coherent signal when your brains communicate through your scalp into the EEG sensor. As a result, it's critical to extract useful information from distorted brain signals and to develop a strong BCI system. When data from electroencephalographic (EEG) sensors is processed through brain-computer interfaces (BCIs), the number of channels, the amount of training data, and the signal-to-noise ratio affect the accuracy of classification [5]. In real-world applications, the SNR is the most difficult to modify of all these variables. The development of several preprocessing and feature engineering methods for reducing noise has been time-consuming and may result in information loss in the extracted features. Feature engineering is heavily reliant on human domain knowledge. Human experience can aid in capturing characteristics of some specific aspects, but it is inadequate in more general situations. An algorithm automatically extracts representative features as the result. Because most machine learning research focuses on static data, it is unable to reliably classify constantly evolving brain signals [6, 7]. BCI systems need to use novel learning methods to deal with dynamic data streams [8]. The benefit of deep learning is backpropagating. By back-propagating, it learns distinguishable information by utilizing the raw brain signals without the need to pre-process and engineer features. Furthermore, deep neural networks may use deep structures to capture latent dependencies and representative high-level features [9].

COMPONENTS OF BRAIN COMPUTER INTERFACE

The primary objective of a Brain-Computer Interface is to detect and analyze the signals in the user's brain that indicate user intention and then transmit these signals to an external computer that executes the user instruction. A BCI-based system has four components that work together to achieve this objective as shown in Fig. (**1**).

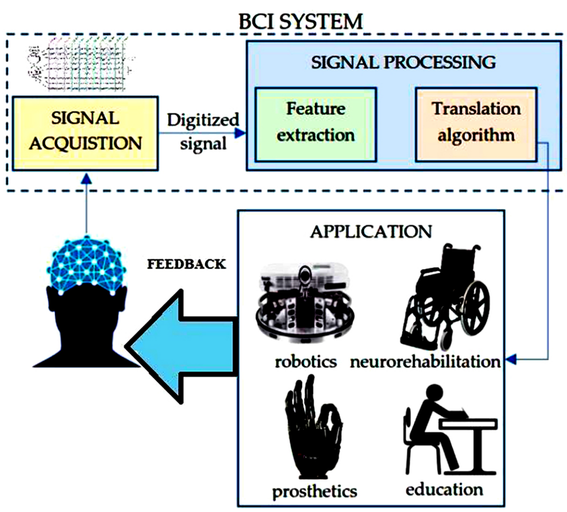

Fig. (1). Components of BCI.

Signal Acquisition

It is the first component that detects and analyses brain signals. The purpose of this component is to receive and register signals produced by neuronal action. It also transfers the resulting signals to the next part of the BCI component for signal enhancement and to reduce the noise [10].

CHAPTER 9

Mining Standardized EHR Data: Exploration, Issues, and Solution

Shivani Batra[1,*], **Vinay Kumar**[1], **Neha Kohli**[2] and **Vaishali Arya**[2]

[1] *KIET Group of Institutions, Delhi-NCR, Ghaziabad, Uttar Pradesh, India*

[2] *GD Goenka University, Gurugram, India*

Abstract: Medical database is among the most crucial databases in terms of their applicability to human life. Many researchers are in search of knowledge that is abstracted within the data. Data mining is popular in today's world as it gives access to knowledge that is otherwise unavailable. The concealed knowledge which is offered as a result of it can help the individual to make better decisions. Data mining tools in health have great potential. These solutions may be divided into four categories: therapeutic efficacy assessment, patient care, customer service, and embezzlement monitoring. The authors discovered that giving decision assistance in the medical sector with an emphasis on electronic health records (EHRs) can save lives. Though offering decision assistance in EHRs using data mining is valuable, it needs consistency. As a result, the authors intend to use data mining methods on standardised EHRs to create a decision support system. This paper presents the state-of-the-art data mining approaches and their application in the healthcare sector. It provides an integrated summary and a comparison detail of the existing literature. This chapter surveys several issues that need to be handled before employing data mining on EHRs and further proposes a solution for dealing with these problems. The problems such as multiple origins, multiple formats, missing data, distinguished users, data granularity, flexibility, and sparseness need immediate attention from researchers. Resolving these problems is important to build an efficient standardized EHRs database.

Keywords: Data Mining, EAV Model, Health Data, Standardized EHRs, Sparseness.

INTRODUCTION

Data mining (DM) is the method of choosing, examining, and modelling massive datasets in order to uncover interesting patterns or correlations that offer the expert a valid and meaningful conclusion [1 - 3].

* **Corresponding author Shivani Batra:** KIET Group of Institutions, Delhi-NCR, Ghaziabad, Uttar Pradesh, India; E-mail: ms.shivani.batra@gmail.com

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

DM is the process of extracting useful information from massive datasets maintained in a variety of places, including data stores, spreadsheets, and cloud platforms. This information is useful in a variety of fields, including corporate strategy, biomedical investigation, and policy decisions. DM is a crucial aspect of information extraction since it examines massive amounts of facts and provides us with previously unrecognized, concealed, and usable information. DM has been successfully employed in a variety of industries, including weather forecasting, healthcare, transit, education, finance, and governance. When employed in a given business, DM offers several benefits such as prediction, clinical diagnosis, and categorization. Aside from such benefits, DM has its own drawbacks, such as the potential for confidentiality breaches. For example, if the miner has access to all of the data's details, he may exploit some of the data's secret information.

Although DM may be used in a variety of fields, its application in the health industry has the potential to assist humanity. Data includes a lot of information that might be concealed. Many academics are working to uncover these previously unknown sections of medical information [4 - 6]. Any platform's most precious asset is data. In the health sector, a lack of vital information can be dangerous to health. The discipline of DM is rapidly expanding, but applying it to medical records is much more difficult than applying it to other data due to the existence of several unique traits.

The remainder of the paper is laid out as follows. Section 2 delves into the complexities of the medical industry, with a focus on EHRs. Section 3 focuses on how DM can be used in EHRs. Section 4 looks at the issues of applying DM to EHRs, and Section 5 offers a remedy. Sections 6 and 7 respectively show related work and the conclusion.

COMPLEXITY IN EHRS

Paper-based patient data have a number of drawbacks, including restoration, productivity, and integrity. EHRs address all of the drawbacks of paper-based health data and make them available to users at a single click.

An Integrated Care EHR [7] is defined as: “a repository of information regarding the health of a subject of care in computer processable form, stored and transmitted securely, and accessible by multiple authorized users. It has a commonly agreed logical information model which is independent of EHR systems. Its primary purpose is the support of continuing, efficient and quality integrated healthcare and it contains information which is retrospective, concurrent and prospective”.

The importance of standardisation for health clients cannot be overstated. Examining the available standards in the health industry is crucial. Various standard bodies are attempting to make interoperable EHRs a reality. Prominent organizations include Health Level 7 (HL7) [8], European Committee of Standardization Technical Committee 251 (CEN TC251) [9], International Standard Organization (ISO) [10, 11], and openEHR [12].

Medical advances are a data-intensive discipline that accumulates massive amounts of complicated and varied information. The health domain is very vast. Considering the complexity of the EHR domain, we identified that it consists of EHR_patients, EHR_concepts, EHR_relationships, terminology concepts and terminology relationships. There are a lot of EHR concepts in a clinical concept. Medical terminologies are a useful tool for standardising taxonomy and incorporating interpretation into EHRs [13]. Professionals frequently use jargon and a plethora of identifiers to define diagnoses. Doctors specialise in a variety of fields. For medical terms, there are several standardised labelling standards. Each specifies its own collection of terminology as well as the connections between them. The SNOMED-CT (Systemized Nomenclature of Medicine-Clinical Terms) coding system, for example, covers about 300,000 medical notions and 7 million connections. Amongst the most significant application fields for DM is healthcare [14]. Apart from having a complex structure, medical data is very sensitive in terms of ethical and legal issues. Besides all of this, a key issue is the occurrence of missing values in the data, which might cause a researcher to receive erroneous findings.

IMPLEMENTING DM ON EHRS

In today's technological age, the volume of health data is growing. Manually processing such a big volume of data is difficult and might result in numerous mistakes. Furthermore, it is likely that a regular person will not be able to acquire hidden knowledge. DM is critical in giving solutions to all of these issues. Many research works are going on to mine electronic health records such as finding the effects of drugs, detecting health insurance fraud, and finding patterns in symptoms to predict future diseases of patients.

All research works which are going on or have already had done provide a good accuracy measure but lack standardization in terms of the local schema used by the researchers. Standardization is necessary since it gives individuals perfect control over any asset, regardless of the criterion. Because of the wide range of storage and retrieval systems, standardisation is critical in the health sector. To construct the data structure, each company has its own set of regulations. This study's major aim is to discover gaps in the use of DM techniques to standardised

CHAPTER 10

Role of Database in Epidemiological Situation

Kanika Soni[1], **Shelly Sachdeva**[1] and **Shivani Batra**[2,*]

[1] *National Institute of Technology, Delhi, India*

[2] *KIET Group of Institutions, Delhi-NCR, Ghaziabad, Uttar Pradesh, India*

Abstract: In this technological era, the technology of databases is very essential to many aspects of modern life. To give the prospective medical practitioner, the finest in class and most recent medical knowledge, it seems mandatory that education in the health domain be well-integrated with the most recent databases. This is because there is a growing demand for it and there are benefits from the collaboration of health-related issues of the public and database technology. Database technology can help improve health in several ways, including connecting geographically separated health providers and patients, collecting data for research studies like drug and vaccine trials, keeping track of chronic diseases, and guaranteeing that patients follow their prescribed treatments. In this pandemic situation of COVID-19, which the whole world is currently suffering, the current paper attempts to emphasize the databases' role. It illustrates how the COVID-19 Dataset can be stored, queried, and analyzed, and helps in providing decision support to various end-users. We have performed descriptive analysis by executing specific queries on the COVID-19 Dataset. Then, we performed predictive analysis using two data analysis techniques on the COVID-19 Dataset to approximate the situation in some major cities of India. Further, we have visualized our results to get valuable information from our analysis.

Keywords: Big data, COVID-19, Data analysis, Data visualization, Database usability, Epidemiological queries, Epidemiological situation.

INTRODUCTION

The technology of databases has advanced significantly over the past few decades, and we can now perform increasingly more complicated queries on larger data sets relatively effectively. However, if we analyze the methodology in which the information is generated, accessed, altered and shared today, we can identify that a significant portion of the world's data is still outside the database systems and the worst part is that we discover a military of the database administrators and other professionals having technical expertise aiding users to import data into and extract it out of a database because users are themselves unable to interact directly

* **Corresponding author Shivani Batra:** KIET Group of Institutions, Delhi-NCR, Ghaziabad, Uttar Pradesh, India; E-mail: ms.shivani.batra@gmail.com

Geeta Rani, Vijaypal Singh Dhaka & Pradeep Kumar Tiwari (Eds.)

with the database due to a lot of factors as majority of the users are naive and can't understand the database schema efficiently if it is complex. Users' ability to access the web directly has been greatly aided by search engines. As a result, users are now able to get information into and out of the constantly evolving web with speed and efficiency. Contrarily, the database community created rigid, precisely defined, and meticulously planned databases under the assumption that the information would be clear, rigid, and well-structured. As a result, databases are now challenging to design, challenging to modify, and even more challenging to query.

Role of Data

Data plays an important role in understanding and managing an epidemiological situation. The key roles of data in epidemiology are described below:

• Surveillance: Data collection and analysis are crucial for continuously monitoring the spread of diseases, identifying trends of spreading a disease in a community, and detecting outbreaks. This data includes lab reports, other modalities such as X-rays, Ultrasound, CT scan reports, history of hospitalizations, and mortality rates.

• Risk Assessment: Data helps epidemiologists assess the risk factors associated with a disease, such as age, gender, geographic location, medical facilities available, vaccination, population . This information aids in identifying the rate of disease spread and populations prone to diseases. Based on this analysis, health experts can guide the public about ways to minimize the disease spread.

• Contact Tracing: Data related to visitors from one place to another along with their medical history helps the health experts to identify the source of spread and minimize the spread. For example, the same has been done during COVID-19 to notify potentially exposed individuals. In case someone was diagnosed with COVID-19 and met with people in a gathering or personally, then he could notify on a social platform or personal mode of communication to take precautionary measures. This helps in minimizing the chances of disease outbreaks.

• Modeling and Predictions: Epidemiological models utilize data to forecast the future course of an outbreak, estimate disease transmission dynamics, and evaluate the potential impact of interventions. For example, as a part of another research we employed the Susceptible, Exposed, Infected, and Recovered (SEIR) model to simulate the impact of vaccination on disease outbreaks. Such models assist policymakers in hospital resource management and reduce the burden on the health industry. Also, such models reduce the mortality rate by informing people at an early stage and reducing the chances of hospitalization.

The amount of data used for above-stated predictions is context-dependent and may vary based on the nature of the disease, transmission dynamics, model employed, desired outcome, *etc.* For example, deep learning models require a huge amount of data while machine learning algorithms require less amount of data for making predictions. But, the more the data, the higher will be the accuracy and reliability of predictions. The data used for epidemiology prediction may hinder the privacy of users. Thus, it is essential to strike a balance between data quantity and privacy considerations.

Privacy-preserving techniques such as de-identification, aggregation, and anonymization can be applied to protect individual identities while using the data for analysis and decision-making. Also, there is a need for finding the right balance between data utility and privacy protection through careful consideration of legal and ethical frameworks, stakeholder engagement, and a transparent decision-making process.

Role of the Database

Database technology is essential to many aspects of modern life in this technological age. By employing online resources to learn about diseases, their signs, symptoms, preventive actions, and general contact information for professionals who may help when needed, database technology has improved human lives. To give the prospective medical practitioner the finest in class and most recent medical knowledge, it seems quite necessary that medical education be well endowed with the most recent databases. Additionally, the updated revision of medical databases might help medical experts to make rapid and accurate decisions with little risk of inaccuracy in today's fast-paced and tough online world. The epidemiology database is available to assist epidemiologists, researchers of public health, health administrators, policy makers, educators and philanthropists in understanding the distribution of diseases, their indicators and determinants, statistics of morbidity and mortality, disease trends, their causes, aggravating factors, and preventative measures. The use of electronic health records, better systems of laboratory for supporting primary and secondary prevention, collection of data for research work like vaccine and drug trials, improving informatics of medical systems with the help of evidence-based, logical, and affordable medication, and surveilling persistent disease conditions are just a few of the ways that database technology can offer a roadmap for improving health.

Epidemiology

Epidemiology is described as the scientific, methodical study, which is data-driven, of the distribution (pattern, frequency), drivers (risk factors, causes), and

SUBJECT INDEX

D

E

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

Geeta Rani

Prof. Geeta Rani received Ph.D. in computer engineering from the University of Delhi, Delhi, India in 2017. She has teaching experience of more than 15 years. She is the chief editor of one book, has published more than 30 research articles, and more than 6 patents. She has acquired more than 17 copyrights on the software applications and inventions. Her research interests include machine learning, artificial intelligence, image processing and data analytics. Prof. Geeta has been a recipient of the 'Women Scientist-IPR Fellowship' in 2018 from the Department of Science and Technology, India.

Vijaypal Singh Dhaka

Prof. Vijaypal Singh Dhaka received Ph.D. in computer engineering from the Bhimrao Ambedkar University, Agra, India, in 2010. He has more than 18 years of industry and academic experience at various organizations of repute. He has more than 80 publications in journals of great repute and guided 11 research scholars to earn PhD. He has published more than 6 patents and acquired more than 15 copyrights on the software applications and inventions. He received "World Eminence Awards 2017" for leading research contribution in ICT for the Year 2016, at WS-4 in London on 15th Feb 2017.

Pradeep Kumar Tiwari

Prof. Pradeep Kumar Tiwari is an associate professor at Dr. Vishwanath Karad MIT World Peace University Pune. He has teaching experience of more than 10 years. He worked as controller of examination for an online education program at Manipal University Jaipur. He has published more than 60 articles in journals and conferences of repute. His research area is artificial intelligence and cloud computing.