INTERNATIONALE
HOCHSCHULE
FERNSTUDIUM

Data Science and Society

(DLMDSSDSS01)

A Research Essay on

Data Science and Analytics in Healthcare

and A Course of Study on

Data Science Use Cases in Healthcare

Author: Kaushik Puttaswamy

Matriculation Number: 321150196

Tutor’s Name: Schlippe, Dr. Tim

Date: 26 November 2022

Table of Contents

Introduction……………………………………………………………………………………………….1

1. Medical Image Processing……………………………………………………………………………...2

1.1 Types of Medical Image Processing…………………………………………………………..….2

1.2 Methodologies and Frameworks in Medical Imaging………………………………………………3

2. Genetics and Genomics…………………………………………………………………………….....5

2.1 Techniques and Frameworks Used for Use Case Analysis……………………………………….5

2.2 Genetic Risk and Gene Expression Use Case Models…………………………………...…….5

3. Creation of Drugs………………………………………………………………………………………..6

4. Virtual Assistances for Patients and Customer Support……………………………………………7

4.1 Four Intelligent Medical Robots…………………………………………………………………….8

5. Predictive Analytics in Healthcare…………………………………………………………………..9

5.1 Use of Predictive Analytics in Healthcare……………………………………………………...9

5.2 Benefits of Predictive Analytics in Healthcare…………………………………………………...9

5.3 Predictive Modelling Process in Healthcare……………………………………………………..10

5.4 Big Data for Predictive Analytics in Health Care……….…………………………………………10

5.5 Examples for Predictive Analytics in Healthcare………………………………………………..11

6. Monitoring Patient Health…………………………………………………………………………...12

Conclusion………………………………………………………………………………………………15

Bibliography……………………………………………………………………………………...…...16

List of Figures

Name of the figure	Figure No	Page No
Predictive Modelling Process in Healthcare	1	10
Big Data Analytics for Flowing Health Data	2	11
IoT System Components	3	13
IoT with Big Data Analytics	4	14

Table of Abbreviations

iDASH	Integrating Data for Analysis, Anonymization, and Sharing
SQL	Structured Query Language
GUI	Graphical User Interface
DNA	Deoxyribo Nucleic Acid
EHR	Electronic Health Record
MRI	Magnetic Resonance Imaging
CT Scan	Computed Tomography Scan
ICU	Intensive Care Unit
AI	Artificial Intelligence
COPD	Chronic Obstructive Pulmonary Disease
USD	United States Dollar
IoT	Internet of Things
3D	Three Dimensional

Introduction

Today, data science is fast expanding to include all global businesses(Team, 2019). In this work, I will illustrate data science use cases in the healthcare industry. Therefore, we will begin to grasp the underlying concept of data science used in medical and health care.

Medicine and healthcare are an innovative and exciting field in which to use data science solutions. Data analytics is taking the medical field to a whole new level by computerising medical records, making new medicines, and researching genetic diseases, among other things. Healthcare and data science are often related through finance, as business strives to minimise expenditures with the aid of vast volumes of data. The fields of data science and medicine are advancing quickly, and it is essential that they progress together(Top 7 Data Science Use Cases in Healthcare, n.d.).

In this paper, I will discuss a variety of use cases pertaining to the healthcare sector, which includes several important fields such as medical image processing, genetics and genomics, drug creation, virtual patient assistance and customer support, predictive analytics in healthcare and monitoring patient health. Each use case is looked at in detail, including how data science analytical tools and methods are used to deal with a variety of data-driven problems in the real world and figure out the best way to solve them.

Data science can provide actionable insights and facilitate the formulation of strategic health system decisions. It aids in the development of a complete view of patients, consumers, and physicians. Decision-making that is informed by data creates new opportunities for enhancing healthcare quality(Subrahmanya et al., 2021).

In this paper, a few factors (use cases) are described that make data science an imperative necessity for the healthcare sector in the current environment, with the competitive desire for quality information in the health market being the most important factor. To acquire accurate analysis to make well-informed decisions on the health conditions of patients. Therefore, it is demonstrated that gathering patient data through the appropriate channels can help improve the quality of healthcare for customers, from physicians to health insurers to institutions. Furthermore, it is outlined that, with the use of data science in healthcare, diseases can be predicted at an early stage, and even remotely, using innovative appliances powered by machine learning. It is explained how mobile applications and smart gadgets continuously collect data regarding heartbeat rates, blood pressure, and sugar levels, transferring this information to doctors in real-time so they can develop remedies accordingly(Data Science in Healthcare – Applications, Roles and Benefits, n.d.).

Moreover, the following work highlights the data science use cases with the greatest effect and future development potential in medicine and healthcare. In addition, different methodologies, frameworks, and tools that enable the integration of numerous types of data, machine learning applications to make data-driven decisions, medical chatbots for virtual assistance, big data and deep learning for predictive analytics in health care, cloud computing and the Internet of Things for monitoring patient health are considered. However, it attempts to provide a brief overview of the few significant healthcare use cases that have significantly impacted the medical industry, and it does not discuss the various roles of a data scientist in healthcare, the challenges and drawbacks of data management, privacy and security, data retention, etc., or the future of data science in healthcare. In addition, this work does not capture the evidence and real-time outcomes of various technological methods adopted to address distinct healthcare use cases.

Data science contributes to the advancement of healthcare facilities and procedures. It contributes to increased productivity in diagnosis and treatment and improves the efficiency of healthcare systems. The fundamental objectives of the healthcare system are increases the productivity, decrease the likelihood of treatment failure, provide appropriate care on time, avoid unnecessary emergencies caused by a lack of doctors, reduce patients' waiting times etc("Data Science in Healthcare - Use Cases and Applications," 2020).

In this work, the reader can expect to find out about some of the most comprehensive data science use cases in healthcare and medicine, as well as the diverse techniques and methodologies that are considered when dealing with such use cases. In addition, the existing benefits and roles played by data science are described while addressing each use case.

The following are the data science use cases having the greatest effect and the most promise for future growth in medicine and healthcare.

1. Medical Image Processing

Medical imaging is the most important use of data science in the healthcare industry. In this field, a lot of research has been done, with big data analytics in healthcare being one of the most well-known. In addition to identifying disease states, medical imaging gives valuable insights into anatomy and organ function. Moreover, it is applied for organ delineation, lung tumour identification, spinal deformity diagnosis, artery stenosis detection, aneurysm detection, and so forth(Belle et al., 2015).

Medical imaging comprises a broad range of image acquisition techniques that are often used for a variety of healthcare purposes with the ease of data science.

1.1 Types of Medical Image Processing

1.1.1 Magnetic Resonance Imaging (MRI)

Deep learning has been used to speed up MRI where the method aims to reconstruct images from under-sampled raw data by learning the regularisation parameter dynamically. These methods take advantage of the expected anatomical appearance and the learnt artefact patterns. The idea is that, during inference, the model can rebuild under-sampled data much faster than the traditional method(Hammernik et al., 2018).

1.1.2 X-ray

Deep learning has become the best way to analyse images, and it has changed the way medical imaging is done(Litjens et al., 2017). Deep learning approaches provide three-dimensional visualisation of X-ray data using simulated X-ray data to train the neural network. A neural network is a set of algorithms that teach a computer to predict outcomes based on the data it receives(Salles & Laboratory, n.d.). Further, with deep learning for chest X-ray imaging, it is also applicable to other X-ray imaging applications, and it is also used in a very important way to measure bone density(Mu et al., 2021).

1.1.3 CT Scan

Automated brain haemorrhage identification using computed tomography (CT) scans and deep learning. In cases of traumatic brain injury, CT is the method of choice for determining the severity of brain haemorrhage. Further The deep learning algorithm uses 3D context from neighbouring slices to enhance predictions at each slice, and then combines slice-level predictions to offer a CT-level diagnosis(Grewal et al., 2018).

Many more techniques are being created to enhance image quality, extract data from images more rapidly, and give the most precise interpretation. Deep-learning-based algorithms get better at diagnosing problems and suggesting better ways to treat them by learning from previous examples.

1.2 Methodologies and Frameworks in Medical Imaging

1.2.1 Hadoop:

Hadoop, a prominent analytic framework, uses MapReduce. The MapReduce is distributed computing framework is used to accelerate and enable three extensive medical image processing use cases(Markonis et al., 2015) that are discussed as follows:

(i) Parameter Optimization for Lung Texture Segmentation Using Support Vector Machines:

The goal of parameter optimization for support vector machines is to find the parameter values that allow high-resolution computed tomography to classify lung textures with the most accuracy(Depeursinge et al., 2012).

(ii) Content–Based Medical Image Indexing:

Content-based image retrieval is another discipline that combines massive volumes of data with computationally difficult tasks. Therefore, indexing large image files using visual characteristics is computed by MapReduce(Yang et al., 2009).

(iii) Three–Dimensional Directional Wavelet Analysis for Solid Texture Classification:

Riesz wavelets provide a multiscale and multi–orientation steerable filter bank that enables the analysis of local orientations and scale with infinite angular and spectral accuracy(Unser et al., 2011). The framework is accessible in 3D for biological image denoising(Chenouard & Unser, 2011). Further using three-dimensional Riesz wavelets, the overall runtime was reduced from more than 130 hours to less than 6 hours while retaining the original Matlab/Octave code and Hadoop streaming(Markonis et al., 2015).

1.2.2 iDASH

It focuses on algorithms and methods for sharing data in a way that protects privacy in biomedical computing projects (Ohno-Machado et al., 2012).

The biomedical and behavioural use cases that drive iDASH are listed below.

I. Molecular Phenotyping of Kawasaki Disease (KD): In this use case to identify new therapeutic targets and tailor treatment. iDASH's secure cloud is utilized to by enabling translational bioinformatics (Sarkar et al., 2011) infrastructure that integrates data from genotyping, gene expression, and proteomic measurements with demographics, laboratory values, images, therapeutic interventions, and clinical phenotypes. Because KD is extremely uncommon, it is important to aggregate data from as many sources as possible. iDASH will make the data accessible through appropriate distribution policies and extensive annotation, allowing for genuine and meaningful reanalysis of the data.

II. Individualized Intervention to Enhance Physical Activity: When remote sensors are employed without the required infrastructure to develop an interventional device that does real-time behaviour pattern identification using machine learning algorithms(Dasgupta, 2011), there is a substantial risk of privacy violation. Therefore, the iDASH cyber-infrastructure is used to share data in a way that protects privacy.

III. Multi-Institutional Surveillance of Medications: Post-market safety surveillance is essential because monitoring drug safety in a distributed environment of newly approved medications is a complicated task that is compounded by the fast diffusion of new medications across previously unstudied patient groups. A higher rate of patient accrual in this type of environment facilitates the early identification of incident rate increases. In addition, due to the rarity of certain of these occurrences, substantial numbers of patients and statistical process controls led by event rate estimates derived from calibrated prediction models are required to identify them accurately(Jiang et al., 2011). The iDASH tool is used for observational cohort outcome surveillance, and its natural language processing technologies are used to pull results from narrative text(Chapman et al., 2011).

In addition, with the introduction of deep learning tools in data science, it is now possible to detect these minute imperfections in scanned photos. Through image segmentation, it is possible to search scanned images for defects(Team, 2019). Overall, several additional methods and frameworks assist in diverse ways in medical imaging, and the most promising applications of data science attempt to identify cancers, artery stenosis, organ outlining, etc.

2. Genetics and Genomics

Genomic data science is a branch of study that allows researchers to employ advanced computational and statistical techniques to extract the functional information encoded in DNA sequences. These data science methods are used in genomic medicine to help researchers and doctors figure out how changes in DNA affect health and illness in people(Genomic Data Science Fact Sheet, n.d.).

Data science tools enable the integration of many types of data with genomic data in disease research, this results in a deeper understanding of genetic concerns underlying responses to certain drugs and diseases(Top 7 Data Science Use Cases in Healthcare, n.d.).

2.1 Techniques and Frameworks Used for Use Case Analysis

2.1.1 MapReduce: Map and Reduce are the two components of the MapReduce algorithm. The Map phase converts a vast volume of data into (key, value) pairs of tuples, while the Reduce phase combines the result of the Map phase into smaller tuple sets(Yadav, n.d.). MapReduce makes it possible to map genetic sequences and cuts down on the time needed to process data effectively.

2.1.2 SQL: SQL is a relational database language that we use to query genomic databases and retrieve data.

2.1.3 Galaxy: Galaxy is an open-source, GUI-based tool for biomedical research that enables you to do numerous genomic operations.

2.1.4 Bioconductor: Bioconductor is an open-source program that is designed to analyse and interpret genetic data.

2.1.5 Deep Genomics: Deep genomics has a tremendous influence on predicting the molecular impacts of important genetic variation for DNA interpretation.

As a result of the continual relationships between genes and external factors, there are still several use cases to recognise. Predictions of genetic risk and gene expression, as well as other complex problems, are now under investigation. Moreover, genetic risk and gene expression are discussed below.

2.2 Genetic Risk and Gene Expression Use Case Models

2.2.1 Genetic Risk Prediction: Genetic risk prediction employs genetic data to individualise the prediction of the result or impact of exposure to a known dangerous toxin(Nebert et al., 2013). In addition, there are numerous approaches for predicting genetic risk, one of which is the mixed model, which is consistently superior to current state-of-the-art approaches. In the widely used liability-threshold model, the mixed model combines the strengths of fixed-effects models, which estimate and add up the effects of single nucleotide polymorphisms, and random-effects models, which rely mostly on similarities in the whole genomes of different people(“Effective Genetic-Risk Prediction Using Mixed Models,” 2014).

2.2.2 Gene Expression Prediction: Gene expression prediction research has traditionally used convolutional layers as its primary architecture. Convolutional layers can be good at finding patterns, but they cannot always simulate how far apart parts of the input sequence are connected (Avsec et al., 2021). Further to include a much broader DNA background, a new model based on transformers, a standard deep learning architecture in natural language processing. Furthermore transformers excel at understanding lengthy text passages(Network, 2022).

Therefore, it is necessary to gather accurate personal genomic data to have a deeper understanding of human DNA. In addition, enhanced genetic risk prediction will be a significant step towards more personalised care.

3. Creation of Drugs

Traditional drug research and development takes a long time and costs a lot of money. According to estimates, *it takes between 10 and 15 years and 58.8 billion USD to get a drug to market*(Biotech R&D Spend Jumps by More than 15% | Nature Reviews Drug Discovery, n.d.). Further only around 250 of the approximately 10,000 evaluated chemical compounds will proceed to clinical testing. In addition, fewer than ten drugs are routinely evaluated in human clinical trials(Schaduangrat et al., 2020). Furthermore from 1995 to 2007, the Tufts Centre for the Research of Drug Development did a study that showed only 11.83 percent of medications that went through phase 1 clinical trials were later approved for sale(DiMasi et al., 2016). In addition, *the success rate of clinical trial medications was just 9.6% from 2006 to 2015*(Swain, n.d.). The increased expense and high failure rate of this conventional drug research and development process have prompted the use of data science.

In recent years, the expansion of data repositories, such as those including chemical and pharmacological data sets, has substantially expanded the availability of large-scale open data for drug discovery. In addition, these domains get regular additions of data, with some libraries comprising tens of millions of chemicals (e.g., PubChem and ZINC databases)(Villoutreix et al., n.d.). The availability of such extensive data sets has a substantial effect on the drug development process. In addition, this method can assist in solving several unmet requirements in drug discovery and design, such that access to this data may facilitate the quick identification of compounds to verify targets or profile disorders, therefore encouraging the creation of new tools and prediction algorithms.

In the middle of the growth of big data (i.e., omics data) that can be used for computational drug development, data curation and pre-processing by database and repository providers make it possible to ensure the quality of this data. Workflows and pipelines in the form of markup languages, codes, or software tools are important for ensuring the reproducibility of computational study(Schaduangrat et al., 2020). The goal of computational drug discovery is to develop computer model simulations as a physiologically appropriate network, which simplifies the prediction of future events with high precision. It enables for the selection of experiments to be conducted and combines all new data into an ongoing learning loop. Moreover, computational drug discovery enhances the collecting and utilisation of many forms of historical data throughout the drug development process. The combination of genetic studies with databases on drug-protein binding may provide remarkable findings. Additionally, it permits testing of chemical compounds against any potential combination of cell type, genetic mutation, and other conditions(Top 7 Data Science Use Cases in Healthcare, n.d.).

In addition, machine learning techniques provide a collection of tools that can enhance discovery and decision-making for well-defined issues, including substantial amounts of high-quality data. At every step of drug research, there are opportunities to employ machine learning. In clinical trials, examples include the validation of targets, the discovery of prognostic biomarkers, and the analysis of digital pathology data. Several context and method-specific applications have been made, and some of them have led to accurate predictions and insights(Vamathevan et al., 2019).

The machine learning application makes it easier to make decisions based on data, and it could speed up the process and reduce the number of failed attempts in drug research and development.

4. Virtual Assistances for Patients and Customer Support

The rapid development of artificial intelligence technology in the medical industry necessitates optimising the clinical process to create a concept in which patients no longer need to see physicians in person. Using a mobile application to bring the doctor to the patient may provide a more effective solution. These AI-powered mobile applications may provide rudimentary healthcare help, often in the form of chatbots. Numerous medical research studies have used chatbots to replicate a physician's diagnostic skills. It is anticipated that chatbots will be able to offer people medical treatment(S et al., 2018). In addition, medical chatbots can do several things, such as pre-consultation, customer service for medical institutions, mental health consulting, and elderly care companionship.

4.1 Four Intelligent Medical Robots

4.1.1 Pre-Consultation Robots

An intelligent pre-consultation chatbot resembles a robot doctor. The patient may describe his or her symptoms to the app, which can then gather the patient's information and send it to the doctor so that the doctor knows about the patient's condition ahead of time. During the pre-consultation phase, the chatbot system will act like a doctor by asking relevant questions based on the patient's symptoms(Mishra et al., 2018). For instance, it will inquire about the beginning of illness as well as the origin, location, colour, and frequency of the patient's disease. It may also inquire about the patient's medical and allergy history.

4.1.2 Medical Service Chatbots

When a patient becomes sick, they may contact the chatbot, which will ask the user a series of questions about their symptoms to identify the disorder. It provides recommendations on the various symptoms to assist patients in understanding their condition. According to the user's response, the chatbot will recommend a physician to consult in the event of a serious illness. For a correct diagnosis, the algorithm remembers how people answered before and asks more specific questions(Rosruen & Samanchuen, 2018).

4.1.3 Mental Health Consultation Chatbots

The digital health of psychological counselling has emerged as a new development trend in psychological treatment(Zhang[1] & Zheng[2]</p>, 2021). In general, the public's conception of chatbots is confined to asking a question and receiving a passive response. However, the chatbot may also pose questions to consumers depending on the context. This mode is more active, facilitates user interaction, and increases user retention. The interactivity of chatbots exceeds that of standard websites.

Cognitive behaviour therapy and powerful natural language processing technologies have been coupled to assist users in recording their feelings and detecting early indicators of depression. In addition, it may provide users with cognitive behavioural therapy, and as the chat progresses, it can identify trends and recommend strategies to reduce unpleasant feelings and ideas(Chung & Park, 2019).

4.1.4 Elderly Care Companion Robots

Companion robots for the elderly can achieve their most basic goal of enhancing the quality of life for the elderly by providing appropriate physical assistance and an ethical companion. However, aged care companion robots still face significant challenges due to poor consumer utilisation and complicated operation(Haleem et al., 2019). Hopefully, in the future, aged care companion robots will be able to address a variety of issues.

Nowadays, the most popular applications that employ this kind of medical chatbot and virtual assistant for customer support are *Your.MD, Babylon Health, Ada, Healthians*, and so on.

The medical chatbot may potentially be used in a substitute manner, while it is unlikely that AI will completely replace human healthcare workers, it may perform certain tasks more consistently, rapidly, and reliably than humans. Moreover, in this manner the optimal customer service is created to achieve a productive relationship between physicians and computers(Top 7 Data Science Use Cases in Healthcare, n.d.).

5. Predictive Analytics in Healthcare

Predictive analytics is a discipline of data analytics that mainly depends on methods such as modelling, data mining, artificial intelligence, deep learning and machine learning. It is used to examine historical and real-time data to create future predictions(Top 7 Data Science Use Cases in Healthcare, n.d.).

Healthcare data is acquired from administrative and medical records, health surveys, illness and patient registries, claims-based datasets, and EHRs and pertains to the health conditions of a person or a group of individuals(Maloy, n.d.). Healthcare analytics is a tool that everyone in the healthcare business may use to offer higher-quality treatment, including healthcare organisations, hospitals, doctors, physicians, psychologists, pharmacists, pharmaceutical firms, and healthcare stakeholders(Infragistics, 2021).

5.1 Use of Predictive Analytics in Healthcare

The healthcare industry generates a tremendous quantity of data but has difficulty transforming it into actionable insights to enhance patient outcomes. In healthcare, data analytics is meant to be used in every element of patient care and operational management. It is used to investigate strategies for improving patient care, anticipating disease outbreaks, and decreasing treatment costs, among other things(Infragistics, 2021).

5.2 Benefits of Predictive Analytics in Healthcare

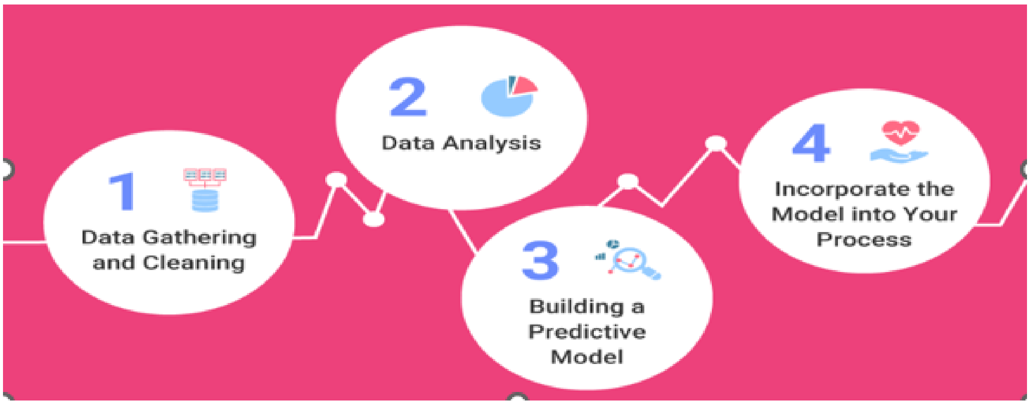
The most important ways predictive analytics can benefit healthcare organisations are(Infragistics, 2021):

a) Overall improved patient care
b) Personalized treatments
c) Population health management
d) Identify at-risk patients
e) Chronic disease management
f) Forecast equipment maintenance needs before they arise
g) Healthcare tracking & digitalization
h) Prevent human errors
i) Fraud detection
j) Reduces overall healthcare costs

5.3 Predictive Modelling Process in Healthcare

The process of predictive modelling involves the execution of prediction-related algorithms on data. Due to the iterative nature of the process, the model that is best suited to the achievement of the goal or business objectives is refined. The process of predictive modelling involves the following steps of analytical modelling(Infragistics, 2021):

Fig 1: Predictive Modelling Process in Healthcare



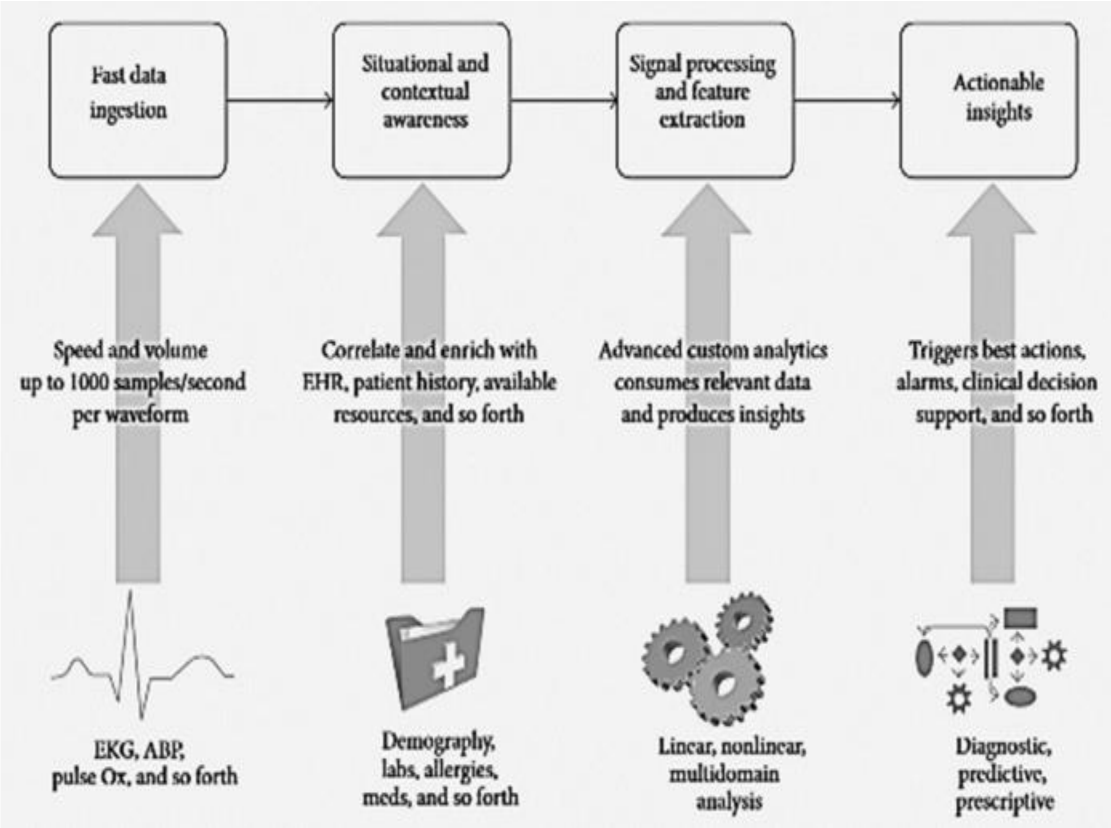
Source: Infragistics, 2021

1) Data gathering and cleansing: It collects data from all sources to extract necessary information by removing noisy data via data cleaning activities, allowing for accurate prediction.
2) Data analysis: Before developing a model, we must understand the behaviour of the data and the relationships between variables.
3) Building a predictive model: While analysing the data, execute as many algorithms as possible and compare their results. Find the test data and use the rules for classification to see how well the classification model works on test data.
4) Incorporate the model into business process: To make the model useful to a healthcare organisation, it must be included into the organization's procedures so that it can be used to enhance patient care.

5.4 Big Data for Predictive Analytics in Health Care

Big data plays a crucial role in predictive analytics, particularly in the medical arena. Big data can process a vast amount of data acquired from the health care sector, which assures answers to the most important problems in the medical industry. Big data analytics in the area of medicine appears as a potential technique for collecting the huge amount of health care data acquired from patients and the public, which can then be used for improved prediction, performance, innovations, and comparative effectiveness. Incorporating big data and next-generation analytics into clinical research, the data reserves become limitless sources of informed information that power the health care system(Smys, 2019).

Fig 2: Big Data Analytics for Flowing Health Data



Source: Krumholz, 2014

5.5 Examples for Predictive Analytics in Healthcare

5.5.1 Reducing Hospital Readmission Rates

The *average readmission rate for adults was 14%, with 20% of readmissions due to one of four conditions: diabetes, heart failure, COPD, and septicaemia*. Using socioeconomic data, EHRs, and predictive analytics, patients with a high risk of readmission may be identified, alerted, and given extra medical treatment to decrease readmission rates(Overview of Clinical Conditions With Frequent and Costly Hospital Readmissions by Payer, 2018 #278, n.d.).

5.5.2 Research into New Treatments

Predictive analytics may also be used efficiently in the study of new medicines. The predictive algorithms may effectively anticipate the individual's reaction to a medicine or therapy by analysing genetic information, clinical history, and other data. This may facilitate the study process and reduce the requirement for inpatient groups(Infragistics, 2021).

5.5.3 Health Insurance

In healthcare, predictive analytics is the ability to compute the precise cost of health insurance for each person based on age, gender, medical history, insurance case history, heredity, etc(Infragistics, 2021). Moreover, predictive analytics may be used to avoid false insurance claims(Predictive Analytics For Insurance Fraud Detection - Wipro, n.d.).

5.5.4 Detecting Early Signs of Patient Deterioration in ICU and General Ward

Patients with the greatest likelihood of requiring an intervention within the next 60 minutes may be identified using predictive algorithms. This lets caregivers act early and proactively if there are small signs that a patient's health is getting worse(Predictive Analytics in Healthcare, n.d.).

5.5.5 Delivering Predictive Care for At-Risk Patients in Their Homes

Predictive analytics may incorporate data from many sources, including hospital-based electronic medical records, fall detection pendants, and previous usage of medical alert systems, to identify seniors who are at risk for emergency transport within the next 30 days(Delivering Predictive Care for At-Risk Patients in Their Homes, n.d.). This lets health care workers reach out to seniors before they have a fall or other health problem.

5.5.6 Identifying Equipment Maintenance Needs Before They Arise

Using predictive analytics, sensors in an MRI scanner may provide technical data for proactive remote monitoring and analysis, highlighting imminent technical faults for prompt replacement or repair(The Rise of the Digital Twin, n.d.).

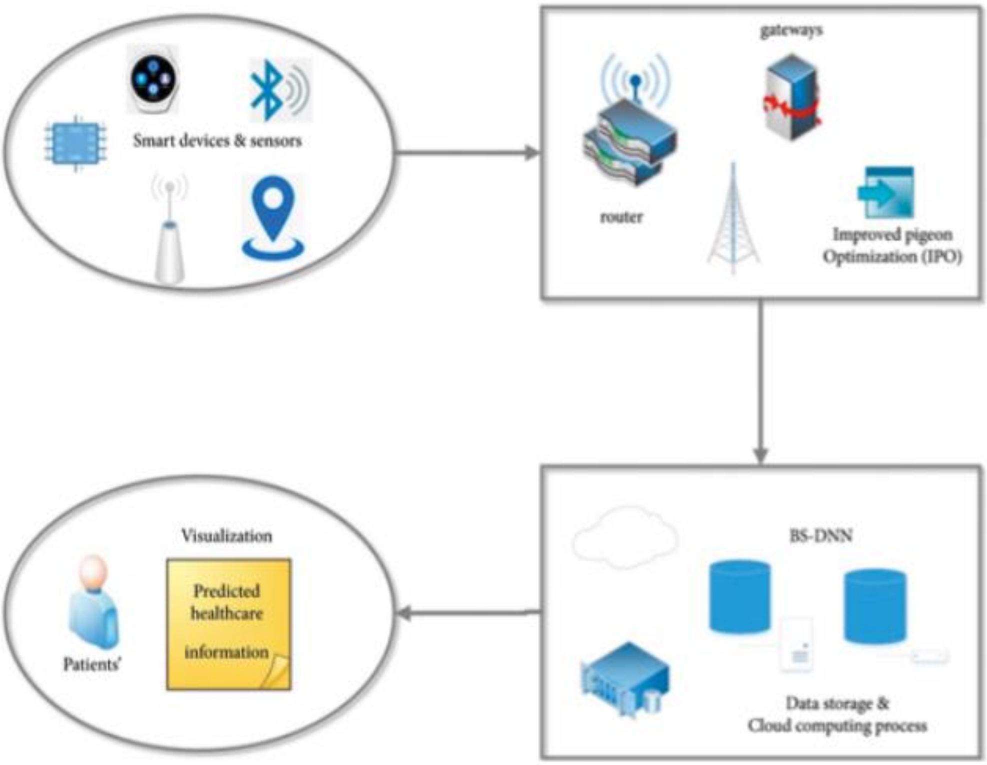
Overall, predictive analytics provides healthcare organisations with a real-time, contextual view of their data, enabling healthcare professionals to provide better care by empowering them to make more intelligent, data-driven decisions.

6. Monitoring Patient Health

Healthcare monitoring and diagnostics have become an integral aspect of the sector. Many people don't like going to the hospital because they don't have enough time, which can lead to a number of health problems(Chaudhury et al., 2017). Early diagnosis and health prediction are vital to the treatment of many illnesses. With an intelligent healthcare system, the healthcare challenges are addressed. In the current intelligent healthcare system, people anticipate inexpensive support for healthcare systems. The innovation of the Internet of Things (IoT), data science, big data analytics, and cloud computing approaches meets this expectation(Shaikh & Chitre, 2017). Data collection and data storage are two of the most important issues confronting the healthcare industry today. Data analytics and data collection play an important role in the health screening of patients. So, cloud computing and data science are the backbones of every healthcare system in order to deal with the different technological problems that can arise (Xu et al., 2017).

The IoT features several healthcare initiatives that benefit patients, hospitals, doctors, families, insurance companies, and other healthcare professionals. The IoT could be used to track how well patients follow their treatment plans or to respond to any urgent medical needs(Banka et al., 2018). The IoT benefits the healthcare industry by fostering strong relationships between patients and healthcare providers. IoT in modern healthcare requires a redesign that compromises social, economic, and technical factors(Islam et al., 2015). Because of this, IoT has gotten a lot of attention over the past few years, which has led to big improvements in the way healthcare monitoring works. The IoT in healthcare employs wearable sensors that collect a tremendous quantity of data, and cloud computing is the most effective method for dealing with this large amount of data(M Abd El-Aziz et al., 2022). As illustrated, the collected data is also evaluated using cloud computing.

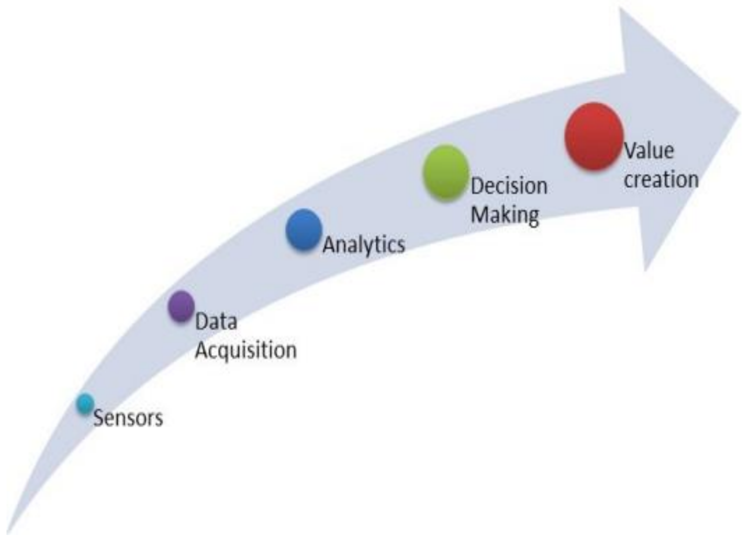
Fig 3: IOT System Components



Source: M Abd El-Aziz et al., 2022

Cloud computing and the IoT are becoming more popular, which is making operations run more smoothly, making staff happier, and making sure patients are safe(Dash et al., 2019). The customised healthcare monitoring system provides e-health services to meet the medical and assistive needs of ageing individuals. The IoT is a remarkable innovation that helps many real-time engineering endeavours with improved administrations. With the use of IoT, it is possible to discover new information with an early forecast, and this information aids in decision-making for better life quality(Jagadeeswari et al., 2018). The Internet of wellness sensor objects are a component of a healthcare monitoring architecture based on the IoT. These gadgets create enormous quantities of data that a physician would find impossible to process. The fundamental concern of the physician is that he must make judgments on a patient's healthcare, which requires separating information about a single patient from a flood of medical data related to a huge number of patients. In this instance, a controller of the IoT would be used to transport medical data to the cloud, which would manage the huge amount of data and enable big data analysis. This data processing and analytics allow for continuous patient condition monitoring. Moreover, the incorporation of the IoT with big data analytics provides real-time monitoring of the acquired information and allows the timely adoption of optimal actions to preserve consistency and throughput. In the case of health care, it provides clinicians with real-time monitoring of patient health and enables them to choose the ideal treatment option in a timely manner. This will enable a reduction in both the treatment's duration and expense(Chaudhury et al., 2017). The figure below depicts the combination of big data analytics with the IoT.

Fig 4: IOT with Big Data Analytics



Source: Chaudhury et al., 2017

Therefore, the growth of big data analytics, which includes the IoT and cloud computing, helps people make better decisions and plan, coordinate, lead, command, and manage things better.

Conclusion

The data science use case solutions in the medical industry have been altering the industry with their multiplicity of tools and techniques (Artificial Intelligence, Machine Learning, Big Data, Deep Learning, etc.) that find new insights and make bold ideas a reality. The opportunities for merging data science and healthcare are vastly expanding as the volume of data grows exponentially faster every day and technology continues to evolve.

In this work, I have discussed a prospective use case that is now impacting numerous healthcare industries, such as medical image processing, where various image acquisition techniques enable clinicians to detect small tumours that would otherwise be difficult to detect.

Researchers use data science to analyse genetic sequences and look for correlations between the sequence's properties and the disease. In addition, genomics research involves locating the optimal medicine, which provides a deeper understanding of how a drug interacts with a specific genetic condition. With the growth of data science-enabled drug discovery, it is now possible to enhance the collection of historical data to aid in the drug development process. Combining genetics and drug-protein binding databases allows for the development of new discoveries in this field.

Data scientists have created a comprehensive virtual platform that aids patients. Using these platforms, a patient can input his or her symptoms and receive information on the numerous possible diseases based on the confidence rate. Using data science, hospitals may be able to predict when a patient's health will start to get worse and take preventative steps or start treatment early to make it less likely that the patient's health will get worse even more.

Furthermore, leveraging IoT devices, such as wearable gadgets that monitor the heartbeat, temperature, and other medical data of the user. The acquired data is analysed with the help of data science. In addition, data science is employed to integrate the numerous sources of industrial knowledge that will be utilised jointly in the treatment process. Through data science, healthcare firms are optimising inventory management, supplier risk management, and other supply chain activities.

Moreover, the number of use cases is continuously expanding, and many general use cases, such as fraud detection and robotization, are applied to healthcare. Nevertheless, some of them are exclusive to the medical industry.

Overall, data science has significantly transformed healthcare and the medical industry, and the future appears bright and optimistic.

Bibliography

Avsec, Ž., Agarwal, V., Visentin, D., Ledsam, J. R., Grabska-Barwinska, A., Taylor, K. R., Assael, Y., Jumper, J., Kohli, P., & Kelley, D. R. (2021). Effective gene expression prediction from sequence by integrating long-range interactions. *Nature Methods*, *18*(10), Article 10. https://doi.org/10.1038/s41592-021-01252-x

Banka, S., Madan, I., & Saranya, S. S. (2018). *Smart Healthcare Monitoring using IoT*. *13*(15), 6.

Belle, A., Thiagarajan, R., Soroushmehr, S. M. R., Navidi, F., Beard, D. A., & Najarian, K. (2015). Big Data Analytics in Healthcare. *BioMed Research International*, *2015*, e370194. https://doi.org/10.1155/2015/370194

Biotech R&D spend jumps by more than 15% | Nature Reviews Drug Discovery. (n.d.). Retrieved October 26, 2022, from https://www.nature.com/articles/nrd.2016.135

Chapman, B. E., Lee, S., Kang, H. P., & Chapman, W. W. (2011). Document-Level Classification of CT Pulmonary Angiography Reports based on an Extension of the ConText Algorithm. *Journal of Biomedical Informatics*, *44*(5), 728–737. https://doi.org/10.1016/j.jbi.2011.03.011

Chaudhury, S., Paul, D., Mukherjee, R., & Haldar, S. (2017). Internet of Thing based healthcare monitoring system. *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)*, 346–349. https://doi.org/10.1109/IEMECON.2017.8079620

Chenouard, N., & Unser, M. (2011). 3D steerable wavelets and monogenic analysis for bioimaging. *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2132–2135. https://doi.org/10.1109/ISBI.2011.5872834

Chung, K., & Park, R. C. (2019). Chatbot-based heathcare service with a knowledge base for cloud computing. *Cluster Computing*, *22*(1), 1925–1937. https://doi.org/10.1007/s10586-018-2334-5

Dasgupta, S. (2011). Two faces of active learning. *Theoretical Computer Science*, *412*(19), 1767–1781. https://doi.org/10.1016/j.tcs.2010.12.054

Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: Management, analysis and future prospects. *Journal of Big Data*, *6*(1), 54. https://doi.org/10.1186/s40537-019-0217-0

Data Science in Healthcare – Applications, Roles and Benefits. (n.d.). Retrieved November 6, 2022, from https://www.knowledgehut.com/blog/data-science/data-science-in-healthcare

Data Science in Healthcare—Use Cases and Applications. (2020, May 28). *Intellipaat Blog*. https://intellipaat.com/blog/data-science-applications-in-healthcare/

Delivering predictive care for at-risk patients in their homes. (n.d.). Philips. Retrieved October 30, 2022, from https://www.philips.com/a-w/about/news/archive/case-studies/20180507-delivering-predictive-care-for-at-risk-patients-in-their-homes.html

Depeursinge, A., Foncubierta–Rodriguez, A., de Ville, D. V., & Müller, H. (2012). Multiscale Lung Texture Signature Learning Using the Riesz Transform. *Lecture Notes in Computer Science*, 517–524.

DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, *47*, 20–33. https://doi.org/10.1016/j.jhealeco.2016.01.012

Effective Genetic-Risk Prediction Using Mixed Models. (2014). *The American Journal of Human Genetics*, *95*(4), 383–393. https://doi.org/10.1016/j.ajhg.2014.09.007

Genomic Data Science Fact Sheet. (n.d.). Genome.Gov. Retrieved November 15, 2022, from https://www.genome.gov/about-genomics/fact-sheets/Genomic-Data-Science

Grewal, M., Srivastava, M. M., Kumar, P., & Varadarajan, S. (2018). RADnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 281–284. https://doi.org/10.1109/ISBI.2018.8363574

Haleem, A., Javaid, M., & Khan, I. H. (2019). Current status and applications of Artificial Intelligence (AI) in medical field: An overview. *Current Medicine Research and Practice*, *9*(6), 231–237. https://doi.org/10.1016/j.cmrp.2019.11.005

Hammernik, K., Klatzer, T., Kobler, E., Recht, M. P., Sodickson, D. K., Pock, T., & Knoll, F. (2018). Learning a Variational Network for Reconstruction of Accelerated MRI Data. *Magnetic Resonance in Medicine*, *79*(6), 3055–3071. https://doi.org/10.1002/mrm.26977

Infragistics. (2021, December 13). *Predictive Analytics In Healthcare | Reveal*. Reveal Embedded Analytics. https://www.revealbi.io/blog/predictive-analytics-in-healthcare

Islam, S. M. R., Kwak, D., Kabir, MD. H., Hossain, M., & Kwak, K.-S. (2015). The Internet of Things for Health Care: A Comprehensive Survey. *IEEE Access*, *3*, 678–708. https://doi.org/10.1109/ACCESS.2015.2437951

Jagadeeswari, V., Subramaniyaswamy, V., Logesh, R., & Vijayakumar, V. (2018). A study on medical Internet of Things and Big Data in personalized healthcare system. *Health Information Science and Systems*, *6*(1), 14. https://doi.org/10.1007/s13755-018-0049-x

Jiang, X., Osl, M., Kim, J., & Ohno-Machado, L. (2011). Smooth Isotonic Regression: A New Method to Calibrate Predictive Models. *AMIA Summits on Translational Science Proceedings*, *2011*, 16–20.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005

M Abd El-Aziz, R., Alanazi, R., R Shahin, O., Elhadad, A., Abozeid, A., I Taloba, A., & Alshalabi, R. (2022). An Effective Data Science Technique for IoT-Assisted Healthcare Monitoring System with a Rapid Adoption of Cloud Computing. *Computational Intelligence and Neuroscience*, *2022*, 7425846. https://doi.org/10.1155/2022/7425846

Maloy, C. (n.d.). *Library Guides: Data Resources in the Health Sciences: Clinical Data*. Retrieved November 13, 2022, from https://guides.lib.uw.edu/hsl/data/findclin

Markonis, D., Schaer, R., Eggel, I., Müller, H., & Depeursinge, A. (2015). *Using MapReduce for Large-scale Medical Image Analysis* (arXiv:1510.06937). arXiv. http://arxiv.org/abs/1510.06937

Mishra, S. K., Bharti, D., & Mishra, N. (2018). Dr. Vdoc: A Medical Chatbot that Acts as a Virtual Doctor. *Research & Reviews: Journal of Medical Science and Technology*, *6*(3), 16–20. https://doi.org/10.37591/rrjomst.v6i3.30

Mu, S., Wang, J., & Gong, S. (2021). Application of Medical Imaging Based on Deep Learning in the Treatment of Lumbar Degenerative Diseases and Osteoporosis with Bone Cement Screws. *Computational and Mathematical Methods in Medicine*, *2021*, 2638495. https://doi.org/10.1155/2021/2638495

Nebert, D. W., Zhang, G., & Vesell, E. S. (2013). Genetic Risk Prediction: Individualized Variability in Susceptibility to Toxicants. *Annual Review of Pharmacology and Toxicology*, *53*(1), 355–375. https://doi.org/10.1146/annurev-pharmtox-011112-140241

Network, F. (2022, February 15). *Deep learning and genomics: Predicting gene expression from DNA sequence*. FEBS Network. http://network.febs.org/posts/deep-learning-and-genomics-predicting-gene-expression-from-dna-sequence

Ohno-Machado, L., Bafna, V., Boxwala, A. A., Chapman, B. E., Chapman, W. W., Chaudhuri, K., Day, M. E., Farcas, C., Heintzman, N. D., Jiang, X., Kim, H., Kim, J., Matheny, M. E., Resnic, F. S., & Vinterbo, S. A. (2012). iDASH: Integrating data for analysis, anonymization, and sharing. *Journal of the American Medical Informatics Association : JAMIA*, *19*(2), 196–201. https://doi.org/10.1136/amiajnl-2011-000538

Overview of Clinical Conditions With Frequent and Costly Hospital Readmissions by Payer, 2018 #278. (n.d.). Retrieved October 30, 2022, from https://hcup-us.ahrq.gov/reports/statbriefs/sb278-Conditions-Frequent-Readmissions-By-Payer-2018.jsp

Predictive Analytics For Insurance Fraud Detection—Wipro. (n.d.). Retrieved November 17, 2022, from https://www.wipro.com/analytics/comparative-analysis-of-machine-learning-techniques-for-detectin/

Predictive analytics in healthcare: Three real-world examples. (n.d.). Philips. Retrieved October 30, 2022, from https://www.philips.com/a-w/about/news/archive/features/20200604-predictive-analytics-in-healthcare-three-real-world-examples.html

Rosruen, N., & Samanchuen, T. (2018). *Chatbot Utilization for Medical Consultant System*. 1–5. https://doi.org/10.1109/TIMES-iCON.2018.8621678

S, D., V, I., S, I., M, P., & S, K. D. (2018). A Self-Diagnosis Medical Chatbot Using Artificial Intelligence. *Journal of Web Development and Web Designing*, *3*(1,2), Article 1,2. http://matjournals.in/index.php/JoWDWD/article/view/2334

Salles, A., & Laboratory, A. N. (n.d.). *Now in 3D: Deep learning techniques help visualize X-ray data in three dimensions*. Retrieved October 24, 2022, from https://phys.org/news/2021-07-3d-deep-techniques-visualize-x-ray.html

Sarkar, I. N., Butte, A. J., Lussier, Y. A., Tarczy-Hornoch, P., & Ohno-Machado, L. (2011). Translational bioinformatics: Linking knowledge across biological and clinical realms. *Journal of the American Medical Informatics Association : JAMIA*, *18*(4), 354–357. https://doi.org/10.1136/amiajnl-2011-000245

Schaduangrat, N., Lampa, S., Simeon, S., Gleeson, M. P., Spjuth, O., & Nantasenamat, C. (2020). Towards reproducible computational drug discovery. *Journal of Cheminformatics*, *12*(1), 9. https://doi.org/10.1186/s13321-020-0408-x

Shaikh, S., & Chitre, V. (2017). Healthcare monitoring system using IoT. *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, 374–377. https://doi.org/10.1109/ICOEI.2017.8300952

Smys, D. S. (2019). *SURVEY ON ACCURACY OF PREDICTIVE BIG DATA ANALYTICS IN HEALTHCARE*. *02*, 10.

Subrahmanya, S., Shetty, D., Patil, D., Hameed, B., Paul, R., Smriti, K., Naik, N., & Somani, B. (2021). The role of data science in healthcare advancements: Applications, benefits, and future prospects. *Irish Journal of Medical Science (1971 -)*, *191*. https://doi.org/10.1007/s11845-021-02730-z

Swain, C. (n.d.). *BIO Releases Clinical Development Success Rates 2006-2015 | News*. Cambridge MedChem Consulting. Retrieved October 26, 2022, from https://www.cambridgemedchemconsulting.com/news/index_files/https:/www.cambridgemedchemconsulting.com/news/index_files/8870dadd6004f660df0f3f461d60ebc8-221.html

Team, D. (2019, April 30). *Data Science in Healthcare—7 Applications No one will Tell You*. Data-Flair. https://data-flair.training/blogs/data-science-in-healthcare/

The rise of the digital twin: How healthcare can benefit. (n.d.). Philips. Retrieved October 30, 2022, from https://www.philips.com/a-w/about/news/archive/blogs/innovation-matters/20180830-the-rise-of-the-digital-twin-how-healthcare-can-benefit.html

Top 7 Data Science Use Cases in Healthcare. (n.d.). ActiveWizards: Data Science and Engineering Lab. Retrieved November 1, 2022, from https://activewizards.com/blog/top-7-data-science-use-cases-in-healthcare/

Unser, M., Chenouard, N., & Van De Ville, D. (2011). Steerable Pyramids and Tight Wavelet Frames in $L_2(\BBR^d)$. *IEEE Transactions on Image Processing*, *20*(10), 2705–2721. https://doi.org/10.1109/TIP.2011.2138147

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews. Drug Discovery*, *18*(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5

Villoutreix, B. O., Renault, N., Lagorce, D., Sperandio, O., Montes, M., & Miteva, M. A. (n.d.). Free Resources to Assist Structure-Based Virtual Ligand Screening Experiments. *Current Protein & Peptide Science*, *8*(4), 381–411.

Xu, B., Xu, L., Cai, H., Jiang, L., Luo, Y., & Gu, Y. (2017). The design of an m-Health monitoring system based on a cloud computing platform. *Enterprise Information Systems*, *11*(1), 17–36. https://doi.org/10.1080/17517575.2015.1053416

Yadav, R. (n.d.). *Genetic Algorithms using Hadoop MapReduce*. 48.

Yang, Z., Kamata, S., & Ahrary, A. (2009). NIR: Content based image retrieval on cloud computing. *2009 IEEE International Conference on Intelligent Computing and Intelligent Systems*, *3*, 556–559. https://doi.org/10.1109/ICICISYS.2009.5358101

Zhang[1], <p>Huiying, & Zheng[2]</p>, J. (2021). The Application Analysis of Medical Chatbots and Virtual Assistant. *The Frontiers of Society, Science and Technology*, *3*(2). https://doi.org/10.25236/FSST.2021.030202