Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 11, No. 1, November : 2022

# A REVIEW ON DATA SCIENCE IN HEALTHCARE

Suman Rani
Assistant Professor
Computer Science Engineering
Arya Institute of Engineering and Technology, Jaipur, Rajasthan

Sweta Sharma
Assistant Professor
Dept. of AIDS
Arya Institute of Engineering and Technology, Jaipur, Rajasthan

Arvind jhajhara
Science Student
Adarsh Vidhya Mandir Secondary School, Nagaur, Rajasthan

Rizwanuddin
Science Student
Smart Move Academy, Gopalganj

## Abstract
Clinical innovation, data frameworks, electronic clinical records, wearable and smart devices, and handheld devices all help to carefully alter medical care frameworks. Information science in medical services is being considered by scientists as a result of the industry's ongoing expansion of information digitization. Numerous studies have been conducted regarding the application of information science in the healthcare industry. Large collections of important data on persistent demography, treatment plans, the outcomes of clinical examinations, protection, and other topics are produced by the medical services sector. Information researchers are interested in the data collected from Web of Things (IoT) devices. Medical care frameworks generate enormous volumes of divided, ordered, and unstructured data, which information science aids in processing, making due with, analyzing, and acclimating. This information requires powerful administration and investigation to gain authentic outcomes. The course of information purging, information mining, information planning, and information examination utilized in medical care applications is evaluated and talked about in the article.

## Keywords

Data Science, Big Data, Healthcare, Recommendation Systems, Data Science Algorithms, and Data Analytics.

## I. Introduction-

One of the largest and fastest-growing industries in the world, the medical services sector has recently overcome enormous obstacles to continue growing. Clinical data frameworks, electronic clinical records, wearable and intelligent devices, and handheld devices are mechanically upgraded, changing medical care frameworks generally and cautiously. Along with the advancement of computing processes in the field of medical services, the growth of clinical big


http://doi.org/10.36893/IEJ.2022.V51I11.124-131

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 11, No. 1, November : 2022

data has given scientists and specialists the ability to extract and visualize massive data in new domains.

Large-scale information analysis has several beneficial and life-saving effects when used in medical services. Massive data refers to the vast quantities of data created by digitizing everything, which are then verified and analyzed by clear advancements.

The knowledge of a population (or of a particular person) could potentially aid in the prevention of epidemics, the treatment of illnesses, the reduction of expenses, and so forth. A key component of wellbeing data science is using modern technology and devices to increase productivity. In this way, cloud processing plays an important role. Another perspective is cloud processing, which refers to storing and retrieving data and projects via the Internet via a network of remote servers assisted by the Internet rather than a local server or a personal computer. That is how PC assets are virtualized. The components as shown in figure are related with a typical healthcare recommender system.

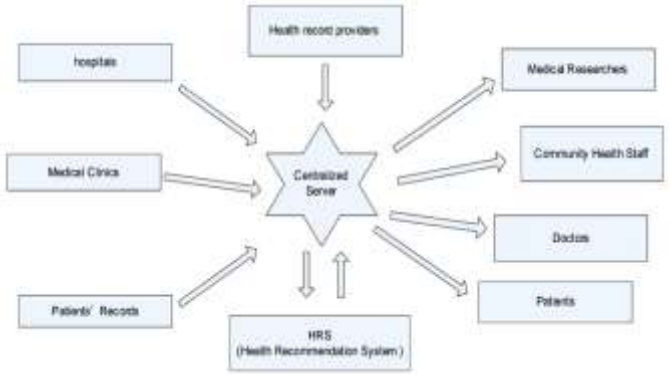


Fig.1 Components of the Healthcare Recommender System

As introduced in [Fig.1] there are numerous partners related to the medical care recommender framework. They include hospitals, medical clinics, patients' records, medical researchers, patients, doctors, and community health staff. A centralized server keeps up with medical care information that is utilized for creating suggestions. The use of huge information examination to acquire a great understanding of finding stowed-away patterns or designs.

## II. Literature Review

Data science is now a major player in healthcare, transforming diagnosis, treatment, patient care, and system effectiveness. The applications, advantages, difficulties, and potential uses of data science in the healthcare industry are examined in this review.

In the context of healthcare, data science uses vast amounts of data to inform choices and improve patient outcomes.

Predictive analytics, machine learning, and natural language processing for clinical decision support are powered by data science, which enhances diagnosis and therapy. Drug development and personalized medicine go hand in hand with telehealth and remote monitoring as they improve patient care.

Precision medicine, lower costs, and better patient outcomes are all produced by data science. It has the potential to save lives by improving outbreak prediction and public health surveillance.

It is necessary to address issues with data privacy, security, and interoperability as well as potential biases and ethical issues. It is essential that the integration with current healthcare systems


http://doi.org/10.36893/IEJ.2022.V51I11.124-131

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 11, No. 1, November : 2022

Blockchain, wearables, wearable data, and the Internet of Things (IoT) have the potential to completely change the healthcare industry. New trends indicate that the industry's future is looking up.

## III. Role Of Data Science in Healthcare

The central job of an information researcher in medical care is to manage huge volumes of information and make significant determinations in light of it. Data Management includes compacting, organizing, and changing information.
Healthcare data scientist usually performs:
Defining the goals of the project as well as the tools and software required
Working with a lot of organized and unstructured information planning to coordinate patient information documents
Cleaning information to meet the association's necessities and targets
Performing information examination for the organisation's inside frameworks and applications
Organizing with engineers to make different models and simulations and register the outcomes

### Use Of Big Data in Healthcare

Big Data has many applications in medical care, changing the business in more than one way. Here are a few manners by which big data is utilized in medical services:

Fig.2 Data Science Use in Healthcare Systems

1) Disease Surveillance and Outbreak Prediction: Big data analytics can assist with observing the spread of sicknesses and foresee episodes. By dissecting information from different sources, for example, electronic wellbeing records, web-based entertainment, and government reports, well-being offices can recognize examples and patterns, empowering them to answer all the more likely plagues.
2) Personalized Medicine: Big data enables the development of personalized treatment plans based on an individual's genetic makeup, medical history, and other factors. This allows for more precise and effective medical interventions, reducing adverse effects and improving patient outcomes.
3) Drug Discovery and Development: Drug organizations utilize big data to examine huge measures of natural and synthetic information to find new medications and improve existing ones. This can speed up the medication advancement process and diminish costs.
4) Predictive Analytics: Medical services suppliers can utilize big data to predict patient results and distinguish those at high gamble of explicit sicknesses. This aids in early mediation, preventive consideration, and asset distribution.


http://doi.org/10.36893/IEJ.2022.V51I11.124-131

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 11, No. 1, November : 2022

5) Clinical Decision Support: Big data can help medical services experts make better choices by giving ongoing data about a patient's condition and therapy choices. It can assist with decreasing clinical mistakes and work on tolerant security.
6) Population Health Management: Medical care associations can utilize big data to analyze the health of
the entire population, distinguish drifts, and allot assets all the more effectively. This can prompt better general well-being drives and the avoidance of ongoing sicknesses.
7) Fraud Detection: Big data examination can assist with distinguishing medical care fraud by recognizing unusual charging designs, overutilization of administrations, and different abnormalities in protection claims and charging information.
8) Remote Patient Monitoring: Big data is utilized to gather and examine information from wearable gadgets and remote sensors, assisting medical care suppliers with observing patients with persistent circumstances and interceding when important.


Fig.3 Application of Big Data in Healthcare

## COMPUTER VISION IN HEALTHCARE

As surprising as it may sound, computer vision goes back as far as the 1960s, when computers started appearing at universities and scientific labs in large number. Healthcare is one of the first industries that recognizes the potential of computer vision. Computer vision is a subfield of "Artificial intelligence". [Fig 4] shows global deep learning in the machine vision deep market.

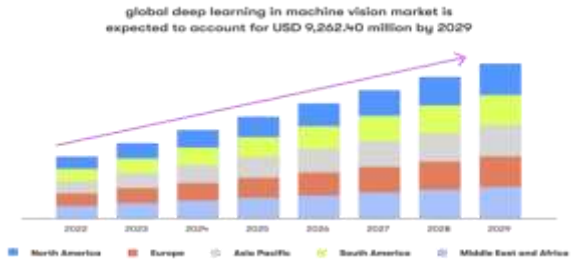


Fig.4 Global Deep Learning in Machine Learning Market

Computer Vision has many applications based on Artificial Intelligence which are shown as follows:
DICOM Image Analysis
Recognition of oddities in X-ray, Feline, and X-beam checks.
Diagnostic assistance


http://doi.org/10.36893/IEJ.2022.V51I11.124-131

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 11, No. 1, November : 2022

Surgical assistance and prevention of inadvertent retention of surgical instruments
Calculation of blood cells
Retina scans and early detection of structural changes

DICOM stands for Digital Imaging and Communications in Medicines, and it is used in image analysis. The exchange and management of medical data and photographs follows an international standard.
Identification of abnormalities in MRI, CAT, and X-ray scans: Medical photos can be obtained in a variety of ways, each with advantages and disadvantages. CT scans use X-rays to produce exact images of the body's internal organs and tissues .
Help with diagnosis: Artificial intelligence (AI)'s machine learning (ML) branch has been effectively used in the medical field to diagnose illnesses. In addition to diagnosing common diseases, machine learning algorithms have demonstrated equal proficiency in diagnosing uncommon diseases.

Surgical assistance and prevention of inadvertent retention of surgical instruments: A comprehensive literature search was performed on MEDLINE®, Embase™, the Science Citation Index, and Google™ Scholar for articles published in English between January 2000 and June 2012. A count of all instruments and sponges should be conducted before, during, and after surgery to ensure that none are left behind.
Calculation of blood cells: We employ a deep learning-based object detection method to detect different blood cells. Among the state-of-the-art object detection algorithms such as regions with convolutional neural network (R-CNN) . [Fig 5] shows the block diagram of the automatic blood cell identification and counting system.

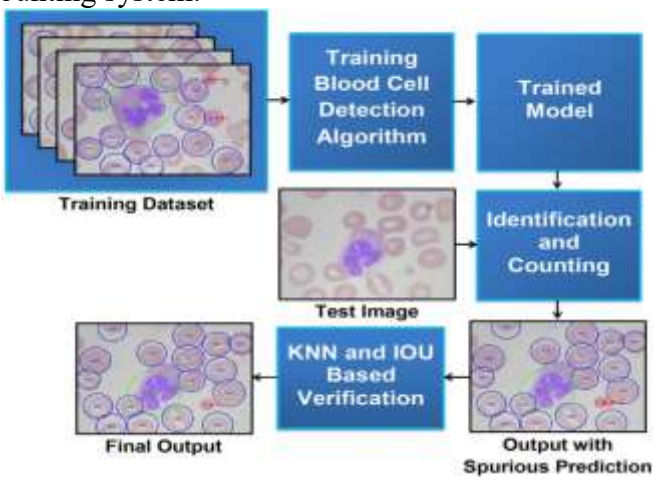


Fig.5 Block Diagram of Automatic Blood Cell Identification and Counting

Retina scans and early detection of structural changes: Retina disease is giving widespread attention because the retina is a leading cause of severe vision loss and blindness. The presentation of intravitreal vascular endothelial development factor (VEGF) hindrance in 2006 massively diminished legitimate visual impairment rates and accomplished impressive enhancements in vision in neovascular AMD and diabetic macular edema (DME).
Privacy and Security in Healthcare:
Data science in the healthcare industry has played a vital role in the privacy and security of medical-related data. The demand had increased to secure big data from hackers. A mobile-based cloud-computing framework of big data has been introduced to overcome the shortcomings of


http://doi.org/10.36893/IEJ.2022.V51I11.124-131

today's medical records systems . We can also encrypt and backup data. Encryption is the process of converting data into code that can only be accessed by an authorized person. Hence, we can encrypt the data and do a backup of it. By utilizing information from gynecology-based reports, Yang et al. framed a system that manually distinguishes characteristics of suspicious specimens from a set of medical care plans that any doctor would mostly adopt.

## IV. Methodologies

Data Science methodologies in healthcare encompass a wide range of techniques and processes to extract valuable insights from healthcare data.

By using these methodologies, we improved decisions, patient care, and advanced medical research. Here are some basic methodologies in data science in the healthcare domain:

Data collection

- Data collection is the foundational step, involving the gathering of various types of healthcare data, including electronic health records (EHRs), medical imaging data, patient-generated data, and genomic data

Data preprocessing:

- Data cleaning is a pivotal move toward the AI (ML) pipeline, as it includes distinguishing and eliminating any missing, copied, or unessential information.
- Data integration: Joining information from numerous sources
- Data transformation: Converting and standardizing data formats and units.

Data Exploration:

- Exploratory Data analysis: Perform EDA to extract valuable insights that help in understanding data's distribution, correlation, and trends.
- Visualization Techniques: Using Scatter plots, histograms, and heat maps to easily understand their patterns.

Feature Engineering:

- Identifying and selecting relevant features for analysis.

Machine Learning and Predictive Modelling:

- It uses different machine learning algorithms, including classification, and regression to make prediction models.
- Basic algorithms used in healthcare include Linear regression, Random forests regressor, decision tree, etc.

Natural Language Processing (NLP):

- NLP strategies are utilized to extricate data from unstructured clinical notes, clinical writing, and patient correspondences.
- NLP helps automate tasks like sentiment analysis, entity recognition, and summarization.

Model Evaluation and Validation:

- Appropriate assessment measurements and cross-approval strategies are fundamental to surveying the performance of predictive models.

Deployment and Integration:

- Successful models are conveyed into medical services frameworks, like EHRs or telehealth stages, to help clinical direction.
- Incorporation with existing medical services frameworks is fundamental for real-world applications.

Privacy and Security:


http://doi.org/10.36893/IEJ.2022.V51I11.124-131

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 51, Issue 11, No. 1, November : 2022

- Medical care information frequently contains delicate data. Information science procedures should address protection and security concerns, utilizing methods like de-identifications, encryption, and access control.
These methodologies are applied in a planned way to saddle the force of information science in medical services, adding to work on understanding consideration, clinical examination, and medical services framework productivity.

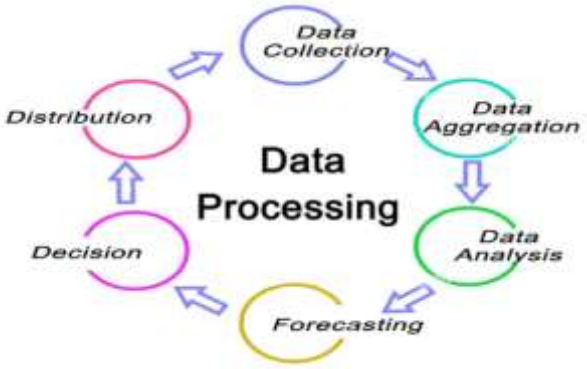


Fig.6 Methodologies of Data Science

## V. Concusion

As we have discussed in the paper, Data science is playing a major role in the healthcare industry in today's life. Big data and machine learning are used to maintain complex medical data. Many AI tools are used in the healthcare industry to treat diseases.
All IT industries are looking forward to improving the IT infrastructure, and data protection and continue working on cloud computing to reduce the cost of all AI tools for storing data.
Technologies such as mobile computing and cloud computing have a great effect in reducing costs and significantly improving the services in the healthcare sector. Data Science can bring in instant predictive analytics that can be used to obtain insights into a variety of disease processes and deliver patient-centric treatment.
Modern healthcare organization is a revolution for medical therapy and medicines by integrating biomedical and health data.


Fig.7 Data Science in Healthcare System

As the medical data is large, it needs proper management and analysis to derive meaningful information.
Predictive accuracy is highly dependent on efficient data integration obtained from different sources to enable it to be generalized.
To summarise, Data Science improves patient care, reduces costs, manages big data, advances in personalized medicines, improvements in machines, and more.


http://doi.org/10.36893/IEJ.2022.V51I11.124-131

Hence, the future of Data Science in the healthcare industry is very bright

## References

[1] "Data Science in Healthcare: A Comprehensive Review" by Shukla, M., & Raghavendra, (2017).
[2] "A Review of Data Science in Health" by Raghupathi, W., & Raghupathi, V. (2014).
[3] "Big Data Analytics in Healthcare: Promise and Potential" by Wang, Y., Kung, L., & Byrd, T. A. (2014).
[4] "Data Science for Healthcare: A Comprehensive Survey" by Choi, E., Schuetz, A., Stewart, W. F., & Sun, J. (2017).
[5] "Big Data in Healthcare: A Review" by Holmes, D. S. (2015).
[6] "The Application of Data Science in Healthcare: A Systematic Literature Review" by Lichtner, V., Dowding, D., Closs, S. J., & Long, A. F. (2018).
[7] "A Comprehensive Review on Data Mining Techniques in the Field of Healthcare Domain" by Rajalakshmi, R., Gopalakrishnan, E. A., & Umamaheswari, K. (2015).
[8] "Data Analytics for Health Care: Concepts, Methods, Tools, and Applications" by Al-Fatih, M. A., Anantharam, P., & Razavi, M. (2019).
[9] "A Review of Data Mining Techniques for Result Prediction in Healthcare" by Maimon, O., & Rokach, L. (2005).
[10] "Data Science in Healthcare: Implications for Risk Management" by Oommen, A., & Prybutok, V. (2018).
[11] "Big Data and Analytics in Healthcare: Introduction and Opportunities" by Chen, H., Chiang, R. H. L., & Storey, V. C. (2012).
[12] Sharma, R., Kaushik, M. and Kumar, G. (2015) “Reliability analysis of an embedded system with multiple vacations and standby”, International Journal of Reliability and Applications, Vol. 16, No. 1, pp. 35-53.
[13] Kaushik, M. and Kumar, G. (2015) “Markovian Reliability Analysis for Software using Error Generation and Imperfect Debugging” , International Multi Conference of Engineers and Computer Scientists 2015, vol. 1, pp. 507-510.
[14] R. Sharma and G. Kumar, "Working vacation queue with K-phases essential service and vacation interruptions," International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014), Jaipur, India, 2014, pp. 1-5, doi: 10.1109/ICRAIE.2014.6909261.


http://doi.org/10.36893/IEJ.2022.V51I11.124-131