ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 1, 2019

Study of Data Science in HealthCare Implications

Dr E. SriDevi[1]

Asst. Professor, Department of C.S.E, Koneru Lakshmaiah Education Foundation, Guntur, A.P, India – 522502.

V.PremaLatha[2]

Assoc. Professor, Department of C.S.E, Koneru Lakshmaiah Education Foundation, Guntur, A.P, India – 522502.

Abstract:

Humans are made by the Almighty with a variety of needs and desires, which links them to their own data. decisions and inclinations. To expand and advance any company or For businesses, it is imperative to understand the needs and desires of their clients.needs of the client based on their data. The changing function of data makes it an extremely important component of any organisation and is carried convinced processes. This paper will present a study that was conducted.concerning Data Science and its applicability to AI,deep learning and machine learning. The application of these data-driven intellectual sciences Science has applications in perming Several operations in our study that we attempted to llustrate the tasks related to data science, such as data cleansing, data processing, and data techniques for modelling, data visualisation, and data presentations. Any business that wants to expand must understand its customers' needs and make wise decisions to meet their expectations going forward creations. The clever data operations or algorithms in the Data science improves the usefulness of data for decision-making and policies for decisions. Also, we emphasise how data science integrates logical, statistical, and mathematical techniques utilising artificial intelligence techniques to reason.We also concentrate on the different data operations tools available in the like SAS, R, Python, and many more.Finally, we turn our attention to how the data science field will fulfil the future demands of numerous businesses. This research paper has the potential to serve as a valuable resource for individuals conducting research and seeking guidance from the data science field when making business-related decisions.

Keyword: Artificial Intelligence (A.I.), Machine Learning (M.L.),Internet of Things (I.OT's) Data Science, Data Analysis, Data Processing, Data Presentations and Data Science Careers.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 1, 2019

1. INTRODUCTION:

Large volumes of both structured and unstructured data are worked with in the study of data science in order to derive actionable insights. This field has the potential to completely transform the healthcare sector through the use of Big Data services, statistical techniques, data mining, and machine learning algorithms. The ability to handle, process, and analyse vast volumes of information produced by hospitals and other medical facilities is made possible by data science. Businesses that want to get invaluable insight into their operational and patient data can use data science consulting services. Healthcare practitioners can utilise these findings to gain a thorough understanding of their patients and to make strategic decisions. In our fast-paced digital age, the intersection of science, technology, and medicine has led to the creation of new data applications that will help advance prescriptive analytics, enhance precision medicine and healthcare personalization, and automate the reporting of health data for clinical decision-making.

There are three ways that data science in healthcare has advanced recently and quickly:

(1) through big data, which is the combination of numerous, complicated data sets, such as digital physiological data from wireless mobile health devices, social media, genetic databases, and electronic medical records;

(2) through brand-new open-access programmes that aim to maximise the data sharing potential of research, clinical trial, and citizen science data sources

(3) in big data analysis methods, such as artificial intelligence and machine learning, which could improve the analysis of both structured and unstructured data.

A number of important questions come to light with the creation, analysis, and increasing availability of new data sets. These questions include the following: How good is the generation of unstructured data? Will data fragmentation and nonreproducible analyses result from using nonstandard methods when processing data using conventional software and hardware? Will big data, particularly from recently released publicly and patient-generated sources, be used and incorporated into healthcare systems? How can big data analytics and newly available open-sourced data benefit medical professionals and researchers. Lastly, how can they learn the necessary abilities to translate knowledge in the data sciences

2. THE ROLE OF A DATA SCIENTIST IN HEALTHCARE

The field of data science holds immense promise for improving the medical domain. Hospitals can use all the patient data they have gathered to improve their diagnostic and

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 1, 2019

preventative capabilities rather than letting it sit idle. In addition, historical data can help medical facilities become more efficient by streamlining the supply chain and inventory management procedures. Working with enormous amounts of data and deriving insightful conclusions from it is the primary responsibility of a data scientist in the healthcare industry. Data conversion, organisation, and compression are all part of data management. The following is a list of typical tasks for a healthcare data scientist:

- Defining the project's objectives and the necessary hardware and software.
- Attempting to organise patient data files by working with a lot of both structured and unstructured data.
- Data cleaning to achieve the goals and requirements of the company.
- Carrying out data analytics for internal systems and apps of the business.
- Collaborating with developers to build different simulations and models and record the outcomes

The primary objective is to arrange the data so that non-technical staff members can easily access and comprehend it. The next stage is to use analytical abilities to identify patterns that can support diagnostic processes. In order to find any correlations or connections, a healthcare data scientist sorts through enormous amounts of data using a variety of models and tests. Afterwards, they offer their observations in dashboards and reports to help doctors make data-driven choices.

BENEFITS OF DATA SCIENCE IN HEALTHCARE

Data science in healthcare raises the standard of public health by offering a host of advantages to patients, doctors, and healthcare and caregiving organisations. Data scientists can contribute to the development of novel drugs, the early detection of epidemics and outbreaks before they become dangerous, and the enhancement of hospital diagnostic capacities through the use of a variety of tools and strategies. We will discuss the main benefits of implementing data science in the medical field.

2. EMERGENCE OF DATA SCIENCE

The term "data science" was first used to describe a new field that dealt with vast amounts of data in the early 1960s. The field of data science is developing, employing statistical methods and computer science to discover new predictions. Data science has

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 1, 2019

advanced significantly to incorporate IoT, machine learning, and artificial intelligence. With the growth of the internet, big data has become a popular concept.

John Turkey discussed the change in statistics in his 1962 paper, The Future of Data Analysis. He emphasised the significance of data analysis in data science in his second paper from 1977. Over the past 30 years, data science has advanced significantly. Today, governments, geneticists, engineers, and other professionals use it globally. Data science is expected to grow significantly in the upcoming years due to a change in how big data is used in the field. Figures 1 and 2 illustrate the expansion of jobs in data science and its evolution, respectively [4].

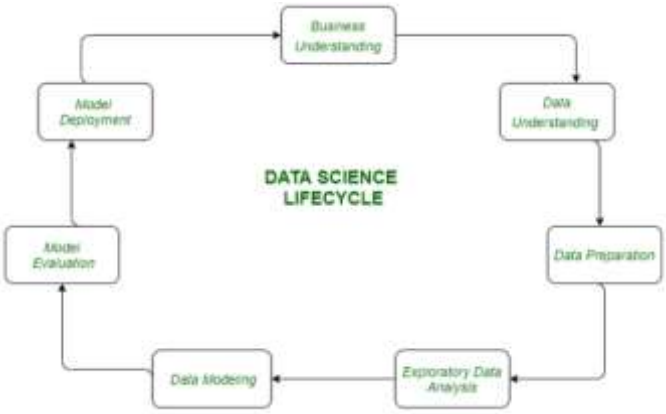

Fig 1 DataScience Life Cycle

3. DATA SCIENCE TOOLS

Data science is one of the most popular fields in the twenty-first century and is growing at an extremely rapid rate. A data scientist bears great responsibility for the analysis and administration of vast amounts of both structured and unstructured data. He needs a wide range of tools and programming languages to work with data science in order to accomplish this. In this paper, some data science tools have been discussed.

The majority of data scientists currently use the following set of data science tools.

I) SAS:

Mostly used for statistical operations, SAS is one of the most popular data science tools. The majority of large organisations use SAS, which is not open-source software, for statistical modelling.For data scientists, this software offers a wide range of statistical libraries and tools. Although SAS has a high degree of reliability, only larger companies use it due to its high cost [5].

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 1, 2019

II) BigML:

Another tool that data scientists frequently use is BigML. It provides a cloud-based graphical user interface (GUI) for processing machine learning algorithms. It specialises in modelling that is predictive.BigML makes use of a wide range of algorithms, including time series forecasting, clustering, and classification [5].

III) MATLAB:

For processing mathematical facts, MATLAB is a multi-paradigm numerical programme. Additionally, the software is not open source.IT makes statistical data modelling possible. Fuzzy logic and neural network simulation are two applications for it. Additionally, it permits signal and image processing. It offers a range of tools for visualisation [5].

IV) Excel:

This data analytics tool is most likely the most used one. Originally created by Microsoft for use in spreadsheet calculations, it is now extensively utilised in data processing, visualisation, and intricate computations. It permits the use of numerous tables, filters, slicers, and formulas. It's also a great tool for cleaning data [5].

V) ggplot2

It is a popular visualisation tool for R programming. It is the library that is most frequently used to create visualisations from data that has been analysed. Even adding text labels and annotations to data visualisations is feasible [5].

VI) Tableau:

Tableau is an interactive visualisation software. The most intriguing aspect of tableau is its ability to connect to databases, OLAP, spreadsheets, and other programmes. Tableau boasts a sizable online community. Tableau Public is the name of Tableau's free edition [5].

VII) TensorFlow

It is a widely used data science tool.software that is frequently used for sophisticated machine learning techniques like deep learning. Since its creation, the open-source software has undergone continuous development. It is compatible with both CPUs and GPUs. In terms of sophisticated machine learning algorithm processing, it is among the best. TensorFlow has several uses, including language generation, image classification, and speech recognition [5].

VIII) Matplotlib:

Matplotlib is a very helpful library that Python has developed for plotting and visualising data. It is widely used to create graphs using the given data. A user can plot or create any

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 1, 2019

complex graph for the provided data using simple lines of code. Data analysts utilise Pyplot, one of the primary Matplotlib modules, to visualise data. Since it is open-source software, anyone can use it, and it is an excellent resource for those just starting to learn data visualisation.

4. DATA SCIENCE APPLICATIONS

Healthcare: Data science can be used to personalise healthcare recommendations and to identify and predict disease.

- Transportation: Real-time shipping route optimisation is possible with data science.
- Sports: Athletes' performances can be precisely assessed by data science.
- Government: Data science can forecast incarceration rates and stop tax evasion.
- Digital ad placement can be automated through data science in e-commerce.
- Data science has the potential to enhance online gaming experiences.
- Social media: Algorithms developed using data science can identify compatible partners.
- Fintech: Credit reports, financial profiles, accelerated underwriting, and predictive models built from past payroll data can all be produced with the aid of data science.

Table 1: Data Science tools and operations

Sl.No	Data Science operations	Tools and operations
1	Data Analysis	Rapidminer, Qlikview, Excel, SAS, Python, Tableau public R and Splunk.
2	Data processing	Hadoop, Cassandra, Cloudera, Flink, Qubole, Statwing, Storm and couchDB.
3	Data presentation	Tableu public
4	Data Scientist role	To breed organizations like medical diagnostics, financial institutions, edification sectors, health care, digital marketing, automated language processing and many more.
5	Future Expectations	Develop natural language processing, business intelligence, social media, whether furcating, stock market predications and others
6	Data scientist professions	Business intelligence developer, data architect, data analysis, data scientist, machine learning scientist and others.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 1, 2019

5. DATA SCIENCE ADVANTAGES AND DISADVANTAGES

The field of data science has benefits and drawbacks of its own. We can now talk about the benefits and drawbacks of data science.

A. Strong Demand

Demand for data science is very high. Social media employment has been expanding quickly, and by 2026, it's expected to generate 11.5 million new jobs.Because of this, it's a highly demanding job in the modern day [11].

B. It's adaptable

Applications for data science are numerous. It is widely utilised in e-commerce, healthcare, banking, and consulting, among other industries. As a result, a data scientist can work in a variety of fields [11].

C. High salary

The field of data science pays well. A data scientist typically makes $116,100 a year. Data science is so popular because of this [11].

D. Enhances Data Quality Data scientists ought to additionally enhance data quality. Their data science focuses on handling, evaluating, and enhancing data quality [11].

E. Produces intelligent goods

Better predictions can be made with data science and machine learning. The development of data science has enabled computers to comprehend human behaviour and make data-driven decisions [11].

F. Has the ability to save lives Data science has greatly enhanced the health care industry. Early-stage tumours can now be identified. Data science has also helped to improve other medical diagnostics, which has improved patient outcomes and saved many lives [11].

DISADVANTAGES:

A. It is impossible to master

Computer science, statistics, and mathematics are all necessary for data science. Consequently, it is impossible to become an expert in every field and to master them all. This is a field that is always evolving [11].

B. Insufficient domain expertise

Domain knowledge is essential to data science. In the field of data science, an individual with limited experience in computer science and statistics will not be successful.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 1, 2019

To succeed in data science, one needs a great deal of experience in this field. This may make switching between domains more challenging [11].

C. Privacy of data

A data scientist must work with a lot of user data. Sensitive user data, such as their citizen ID, may be included in this data. Therefore, there is always a chance that these data won't be kept private. This brings up numerous moral concerns. 11].

D. Fuzzy Word

The term "data science" lacks precise definition and is therefore ambiguous. A data scientist's role varies depending on the field in which they operate[11].

CONCLUSION:

In nearly every aspect, data science has raised living standards considerably. It's getting harder and harder to envision a field without data science's impact. Data science is becoming more and more necessary as the volume of data grows daily. However, there is still a great need for ethically sound data scientists. Data science is expected to significantly raise the calibre of many different domains. This paper discussed the growth, benefits, and range of uses for data science. These days, data science is a required field that bridges several academic disciplines, including mathematics, statistics, logical reasoning, mathematical methods, intelligence algorithms, and machine learning applications. These fields are all related to accessing and effectively using data from different businesses or organisations. When data is used effectively, appropriate decisions can be made to expand the business based on the preferences and satisfaction of the customers. Therefore, we can draw the conclusion that as data science becomes more popular, more data scientists will be needed in order for each organisation to expand. Finally, we turn our attention to the development of successful careers in data science. This field's primary feature is that it once helped all businesses expand.

REFERENCES

1.) Russell, Stuart J., and Peter Norvig. Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited,, 2016.

2.) Nilsson, Nils J. Principles of artificial intelligence. Morgan Kaufmann, 2014. 3.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 1, 2019

3.) Bell, Jason. Machine learning: hands-on for developers and technical professionals. John Wiley & Sons, 2017.

4.) Van Der Aalst, Wil. "Data science in action." Process mining. Springer, Berlin, Heidelberg, 2016. 3-23.

5.) Dhar, Vasant. "Data science and prediction." Communications of the ACM 56.12 (2013): 64-73.

6.) Hazen, Benjamin T., et al. "Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications." International Journal of Production Economics 154 (2014): 72-80.

7.) Wimmer, Hayden, and Loreen Marie Powell. "A comparison of open source tools for data science." Journal of Information Systems Applied Research 9.2 (2016): 4.

8.) Islam, Mohaiminul. "Data Analysis: Types, Process, Methods, Techniques and Tools." International Journal on Data Science and Technology 6.1 (2010): 10.

9.) Nicolae, Bogdan, et al. "Park, Yoonho. Leveraging Adaptive I/O to Optimize Collective Data Shuffling Patterns for Big Data Analytics. IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS. PP (99) pp: 1-13." (2015).

10.) Abas, Zuraida Abal, et al. "Analytics: A Review Of Current Trends, Future Application And Challenges." Journal of Advanced Computer Technology. PP 3560 (2017): 3565.

11.) Rani, Bindu, and Shri Kant. "An Approach Toward Integration of Big Data into Decision Making Process." New Paradigm in Decision Science and Management. Springer, Singapore, 2019. 207-215.

12.) Bejjam, Suvarnamukhi & Seshashayee, M.. (2018). Big Data Concepts and Techniques in Data Processing. International Journal of Computer Sciences and Engineering. 6. 712-714.