Global Journal on Application of Data Science and Internet of Things
http://eurekajournals.com/GJADSIT.html
ISSN: 2581-4370

Revolutionizing Data Science: Recent Advances and Transformative Technologies

Manuj Takkar[1], Preet Deshwal[1], Ojus Jaiswal[1], Yuvraj Singh[1]

[1]Department of Computer Science and Engineering, Chandigarh University, Mohali, Punjab-140413, India.

Abstract

Recent advances in data science have revolutionized the way we collect, analyze, and interpret data. This has been made possible through breakthroughs in machine learning, artificial intelligence, big data analytics, and data visualization techniques. These advancements have enabled us to process large volumes of data quickly and efficiently, uncover hidden patterns and insights, and make data-driven decisions with confidence. Some of the recent developments include the rise of deep learning, natural language processing, and reinforcement learning, which have transformed fields like image recognition, speech recognition, and game playing. Moreover, the advent of cloud computing and distributed computing has made it easier for businesses to leverage data science for their operations. As data science continues to evolve, we can expect further breakthroughs in areas like predictive modeling, causal inference, and automated machine learning. Overall, recent advances in data science have opened up new opportunities for innovation and discovery across industries and domains.

Keywords: Data Science, Artificial Intelligence, NLP, Deep Learning, Applications of Data Science.

Introduction

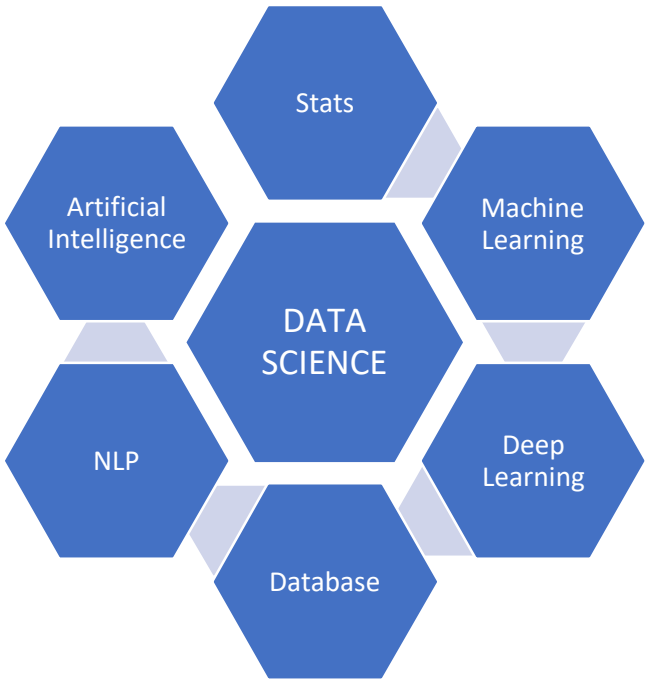
DATA SCIENCE is the application of statistical analysis, machine learning, data mining, and other data- analytic discipliner to large sets of data in order to extract useful information. Data science is not just aboutextracting information from the data. It is also about misunderstanding the meaning of the extracted information and using it for decision making. The recent advancement in data science have helped businesscollect more meaningful insights from their existing datasets and also help them generate new datasets formore accurate predictions. Data Science is a rapidly growing field. It is not just about data and algorithms anymore.

Data Science has been around for decades, but it is only recently that it has started to take the shape of a professional discipline. Data scientists are no longer just statisticians and computer scientists who work ondata projects from time to time. They are now professionals who are

specialized in data analysis and are well informed about the latest developments in this field. The advancement of data science has led to the development of new algorithms and techniques that have enabled us to solve problems that were once considered unsolvable, such as predicting human behavior or detecting cancer at an early stage. Data science is the process of extracting knowledge from data and then converting that knowledge into insights.

Everything in this universe has its own benefits and drawbacks. Similarly, Data Science also has its own benefits and drawbacks. Data Science has made things very easy to collect and analyze data and records but with that so many drawbacks are coming forward with time. We will discuss all this in detail about few of theadvantages and disadvantages of the Data Science.

Parts of Data Science



a) Machine Learning

Machine learning is a set of algorithms that helps computers to learn from data without being explicitly programmed. It is a branch of artificial intelligence, and it can be used to solve problems in many differentfields. Machine learning has been used for various purposes, such as image recognition, predictive analytics, natural language processing (NLP), speech recognition and translation. It has also been applied to the field of medicine with the use of deep neural networks. In this section we are going to explore how machine learning can be applied in data science and what are some use cases where machine learning is applicable. Machine learning is a type of artificial intelligence that provides computers the function with the ability to automatically learn and improve from experience without being explicitly programmed.

Machine learning can be applied for large amounts of data and has many applications in the field of data science. It is a branch of artificial intelligence, which gives machines the ability to automatically learn and improve from experience without being programmed by the user.

Machine Learning is an AI technique thatis based on the idea that data can be analyzed and used to learn from experience. It can help us with a number of tasks, such as recognizing patterns or charts in data, predicting future behavior, and finding newcorrelations. As it becomes more and more popular, ML has been applied to a wide range of industries. This includes finance, healthcare, transportation and retail. In fact, some experts say that ML will eventuallybe able to outperform humans in all occupations because of its ability to process large amounts of data quickly.

The world is changing rapidly as AI systems become smarter and smarter every day.

b) Deep Learning

Deep Learning is a part or subset of machine learning and it is used for making predictions from data. Deep learning is inherited from machine learning and it is used for making predictions from data. It can be used for tasks like image classification, speech recognition, natural language processing (NLP), and more. Deep learning algorithms are usually trained with large datasets which are called supervised deep learning. They can be trained with labeled images or videos that have caption information that tells the algorithm what the contentof the video or image is. The other type of deep learning called unsupervised deep learning doesn't need any labels to train the algorithm because it learns by itself what patterns exist in the data set by looking at similar examples. This type of deep training doesn't require much work but it will take longer time.

Deep learning is a subset of machine learning that uses artificial neural networks to train itself. These networks are composed of many layers, with each layer containing a number of neurons. The first layer is called the input layer, and it receives input from the outside world. The last layer is called the output layer,and it sends information back to the outside world. In between these two layers are a number of hidden layers, which are responsible for processing information in order to produce an output that can be understood by humans. Each neuron in these hidden layers connects to several neurons in previous layers and outputs its own set of values based on those connections. Deep learning has been used for many different applications including facial recognition software, language translation software, self-driving cars,and more.

c) Database

When data is large or commercial data which is used in businesses like banks, retail stores, and manufacturing firms, so they will have a lot of organized data in the form of database tables and these data are stored in the disk, simply database is the collection or arrangement of data in some a model like the relational model in the permanent storage so that it can be accessed applications easily, this arrangement in the hard disk is called a database. We can also say that database is like a container that stored the data in the proper arrangement so that we can modify that data easily when we need it. It helps to solve real-world problems. It also controls and helps in data sharing and increases the security of data.

Data technology is one of the fast-developing fields that I cannot see slowing down anytime soon. Not with how our facts dependence is overgrowing day with the aid of using day. Data technology is all approximately facts, gathering them, cleansing them, studying them, visualizing

them, and the use of themto make our existence better.

A database is described as an established set of statistics held in a computer`s reminiscence or the cloud this is on hand in numerous ways. As a statistics scientist, you may want to design, create, and engage withdatabases on a maximum of the tasks you may paintings on. Sometimes you may want to create the wholething from scratch, even as at different times, you may simply want to recognize the way to talk with an already the current database.

d) Artificial Intelligence

What is Artificial intelligence in terms of data science?

AI is only a pc. This is capable of simulate human concept and behavior. AI is a group of mathematical algorithms that make computer systems apprehend complicated relationships and make actionable decisions, and plan for the future. AI has very important role in the field of Data Science.

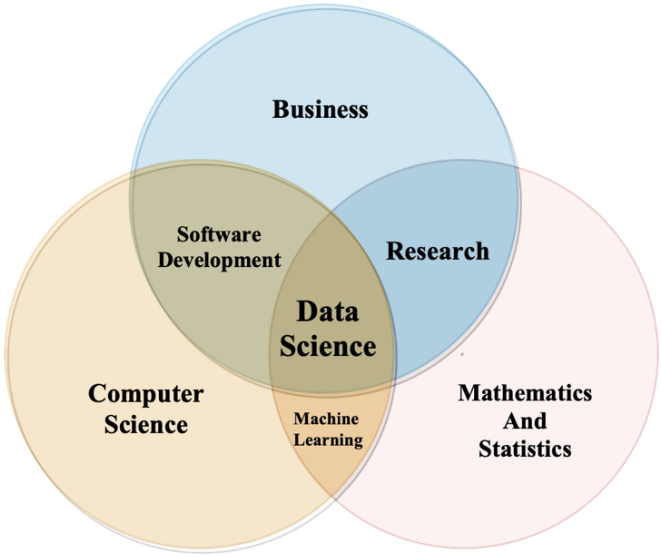
IBM WATSON is an AI generation that facilitates physicians fast become aware of records in a affected person's clinical report to offer powerful remedy to the affected person with the aid of using taking facts then gives its evident-primarily based on personalised advice full-filled with the aid of using records froma series of 350+ journals, 250+ textbooks, and 20+ pages of texts which offer docs immediately get entry to a wealth of records personalised to the affected person's remedy plan.

BLUBERRY: This robotic can perform comedy after being accumulate subtitles from masses of heaps of movies .Kory Mathewson a synthetic intelligence researcher on the University of Alberta, Edmonton, created an set of rules designed to riff with him onstage. He skilled it to create traces of discussion for usein an overall performance with the aid of using worthwhile it whilst the speak makes sense.

AI may be used for production technique improvement, processing biomedical and scientific data, creating“clever assistants” (like google assistant, Alexa, Siri, etc.) or chat bots, social media monitoring, economicplanning, investing, and plenty of different fields.

e) Statistics

Statistics is one of the most important part of data science. It provides methods to find more insights of data. Statistics is an important part of data science because it provides unbiased information about a givensample population. Statistics can be used to determine how confident one can be in the results they have found or if they are statistically significant. Statistics can be broadly defined as the branch of mathematicsdealing with probability, sampling, and inference. The term statistics was introduced by William Playfair in 1786 and was derived from an earlier word statistician which was used for reference to people who computed mathematical tables for summarizing numerical data on economic topics such as national income and expenditure.



f) NLP

NLP stands for Natural Language Processing. Natural Language Processing (NLP) is a field of computer science which deals with the interactions between computers and human (natural) languages. NLP includes the study of natural language processing algorithms, which are often used to understand the meaning of text, speech and other forms of human communication. NLP is a subset of artificial intelligence and machine learning. It is also an important part of speech recognition software, search engine indexing, optical character recognition and other text-processing software applications. NLP is a field of study that is focused on the interactions between computers and human language. It is an important part of data science, but it is also used in other fields, such as computational linguistics and artificial intelligence.

NLP is an important component of many data science applications, especially in the area of machine learning. NLP can be used to parse text, identify sentiment, and extract meaning from unstructured text. NLP is also used in sentiment analysis, which is the process of understanding the sentiment behind a piece of text. Sentiment analysis can be used for customer service and marketing efforts to understand how customers feel about products or services. For example, if a customer complains about their product on social media, it is important for a business to know how the customer feels about their product so they can offer appropriate resolutions such as discounts or coupons.

Applications of Data Science

In this section, we will be exploring the different use cases of data science. We will also explore how it hasbeen applied to various industries such as finance and marketing.

Data science is a field of study that uses data to identify patterns and trends in order to make predictions about future events. Data scientists are the ones that use statistical analysis, machine learning, and other methods to extract insights from datasets.

The applications of data science are vast. It can be used in many diverse fields such as healthcare, finance, marketing and more. For example, it can be used in the healthcare industry to predict when a patient may have an emergency or need hospitalization by analyzing their records and vitals. Data science has becomea popular field in recent years. Its applications are diverse, from social media marketing to financial modelling. The job market is also booming with many data scientists being recruited by major companies and even startups. Data scientists have the potential to make a huge difference in any industry they enter.

Data science is the study of data and its analysis. Data science can be applied in a wide range of fields, from marketing to healthcare. It is a very useful tool for many companies because it analyzes data, identifies patterns and trends and provides insights that can be used in decision-making processes.

Advancements of Data Science

The advancement of Data Science is a double-edged sword. While it helps businesses and organizations make better decisions, it also has the potential to create a dystopian society where personal information canbe used against its owner. Data science is a relatively new field that applies scientific methods to data in order to extract insights and understanding. The term data science is often used interchangeably with the terms big data, machine learning, and artificial intelligence. Data science has been around since the 1980sbut it has become more popular in recent years.

Data science is a field of study that uses data to understand and extract knowledge from it. Data science isa broad term that includes many different academic disciplines that use approaches such as statistics, machine learning, data mining, pattern recognition, and data visualization. Data science in healthcare has helped improve health outcomes for patients by predicting their risk for various diseases and conditions. This has been possible because of the tremendous amount of data that is available in this field. Data sciencehas become an indispensable part of the modern world. It can be found in almost every industry and organization including social media networks, hospitals, governments, research labs etc.

The advancements of data science are changing the way we live our lives by impacting things like healthcare, education, and marketing. Data scientists have created algorithms that can predict when someone will die based on their medical records. They have also created an algorithm that predicts which students will be successful in college based on how they did in high school. This is a powerful tool for colleges to use when deciding who to admit into their programs.

Future Scope of Data Science

Data Science's Contribution to the Future:

Data Science encompasses many leap forward tech principles like Artificial Intelligence, Internet of Things, Deep Learning to call a few. With its development and technological developments, information technology's effect has expanded drastically.

a) VR AND AR

Product advertising and customer support can be revolutionized with the aid of using superior chatbots, Virtual Reality (VR), and Augmented Reality (AR). We is probably searching ahead to a time whilst personalized patron revel in will encompass stay simulations, interactive demos, visualization of proposedsolutions.

b) Health Care Area

There is a massive requirement of information scientists withinside the healthcare area due to the fact they invent quite a few information on a everyday basis. Tackling a large quantity of information isn't alwaysfeasible with the aid of using any unprofessional candidate. Hospitals want to hold a file patients' clinical history, bills workforce non-public history, and plenty different information. Data scientists have become employed withinside the clinicals area to decorate the pleasant and protection of the information.

c) Transport Sector

The delivery area calls for a information scientist to research the information amassed via passenger counting systems, asset management, area system, fare collecting, and ticketing.

d) E-Trade

The E-trade enterprise is booming simply due to information scientists who examine the information and create custom designed advice lists for presenting incredible effects to end users

e) Sales and Marketing

Data evaluation examines records received from advertising campaign to find out developments including how a method affects sales, customer behavior, geographical preferences, and innovating preferences

Conclusion

So, till now we have discussed about What is Data Science? Future scope of data science, Advancements in data science, Application of Data Science. These are the main highlights and points which a person mustbe having knowledge of if he or she wants to work in the field of data science.

Now, its time to conclude and make a short summary which is also a very important part of a paper. In the introduction part we have discussed about what is data science, in that we have conclude the basic knowledge about data science. After that we have discussed about the Parts of data science. Basically, data science has 6 parts i.e., NLP, AI, ML, Deep Learning, Stats, Database. In NLP section, we have discussedabout what is NLP and why it is important and where it is used. After that we have discussed about AI (Artificial Intelligence). What is the role of AI in data science and why it is so important? Then we have discussed about Database, the role of data base in data science and where it is used. Then we discussed about Machine Learning, what is the role of ML in data science and its benefits. The we discussed about Deep

learning, what is the use of deep learning and what is its role in data science. Then we have discussedabout the most important part of data science i.e., Statistics. We have discussed what is Stats, why it is soimportant and what is its role in data science. Then we have discussed about Applications and Future scopeof data science.

Refrences

Dhar V. Data science and prediction. Communications of the ACM. 2013 Dec 1;56(12):64-73.

Provost F, Fawcett T. Data science and its relationship to big data and data-driven decision making.Big data. 2013 Mar 1;1(1):51-9.

Van Der Aalst W. Process mining: data science in action. Heidelberg: Springer; 2016 Apr 15.

Ashabi A, Sahibuddin SB, Haghighi MS. Big data: Current challenges and future scope. In2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE) 2020 Apr 18 (pp. 131- 134). IEEE.

Baesens B. Analytics in a big data world: The essential guide to data science and its applications. JohnWiley & Sons; 2014 May 19.

Gruson D, Helleputte T, Rousseau P, Gruson D. Data science, artificial intelligence, and machine learning: opportunities for laboratory medicine and the value of positive regulation. Clinical biochemistry.2019 Jul 1; 69: 1-7.

Weihs C, Ickstadt K. Data science: the impact of statistics. International Journal of Data Science andAnalytics. 2018 Nov; 6(3): 189-94.

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y. Reproducible, interactive, scalable and extensible microbiome datascience using QIIME 2. Nature biotechnology. 2019 Aug; 37(8): 852-7.

Waller MA, Fawcett SE. Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. Journal of Business Logistics. 2013 Jun; 34(2):77-84.

Cai L, Zhu Y. The challenges of data quality and data quality assessment in the big data era. Datascience journal. 2015 May 22;14.

Cai L, Zhu Y. The challenges of data quality and data quality assessment in the big data era. Data science journal. 2015 May 22;14.

Brunton SL, Kutz JN. Data-driven science and engineering: Machine learning, dynamical systems,and control. Cambridge University Press; 2022 May 5.

Ostle B. Statistics in research. Statistics in research.. 1963(2nd Ed).

Lee YG, Kim SY. Introduction to statistics. Yulgokbooks, Korea. 2008:342-51.

Cumming J, Riggins V, Hodson P, Walker B. Advanced Well Planning Using Natural Language Processing NLP and Data Science Models: Maximizing the Value of Data to Mitigate Costs and Risks in New Wells. InAbu Dhabi International Petroleum Exhibition & Conference 2020 Nov 9. OnePetro.

Zhu L, Zheng WJ. Informatics, data science, and artificial intelligence. Jama. 2018 Sep 18;

320(11):1103-4.

Krishna CV, Rohit HR. A review of Artificial Intelligence methods for data science and data analytics:Applications and Research Challenges. In2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2018 2nd International Conference on 2018 Aug 30 (pp. 591-594). IEEE.

Sharda R, Delen D, Turban E. Analytics, data science, & artificial intelligsssence: Systems for decision support. Pearson Education Limited; 2021.

Baldi P. Deep learning in biomedical data science. Annual review of biomedical data science. 2018Jul 20;1:181-205.

Nosratabadi S, Mosavi A, Duan P, Ghamisi P, Filip F, Band SS, Reuter U, Gama J, Gandomi AH.Data science in economics: comprehensive review of advanced machine learning and deep learning methods. Mathematics. 2020 Oct;8(10):1799.

Kanter JM, Veeramachaneni K. Deep feature synthesis: Towards automating data science endeavors. In2015 IEEE international conference on data science and advanced analytics (DSAA) 2015Oct 19 (pp. 1-10). IEEE.

Beysolow II. Introduction to Deep Learning. InIntroduction to Deep Learning Using R 2017 (pp. 1-9).Apress, Berkeley, CA.

Gutierrez DD. Machine learning and data science: an introduction to statistical learning methods with R. Technics Publications; 2015 Nov 1.

Provost F, Fawcett T. Data Science for Business: What you need to know about data mining and data-analytic thinking. " O'Reilly Media, Inc."; 2013 Jul 27.

McFadden FR, Hoffer JA. Database management. Benjamin-Cummings Publishing Co., Inc.; 1991Jan 1