Journal of Information Systems Engineering and Management
2025, 10(19s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Data Science: A Concept, Applications, Challenges, Technologies and Case Studies for Engineers

K Dhananjaya Kumar[1,2,*], N Vinutha[1,2], M B Meghashree[1,2]
[1]*Computer Science and Engineering, Vidya Vikas Institute of Engineering and Technology, Mysuru-570 028, Karnataka, India*
[2]*Visvesvaraya Technological University, Belagavi-590 018, Karnataka, India*
[*]*Corresponding Author: https://orcid.org/0009-0001-2076-0483*
kdhananjayakumar@gmail.com

ARTICLE INFO

Received: 05 Oct 2024
Revised: 05 Dec 2024
Accepted: 22 Dec 2024

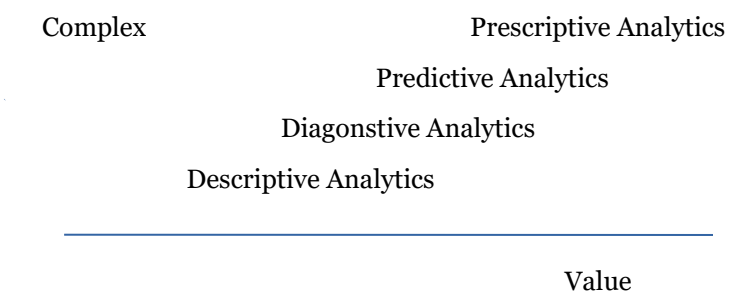
ABSTRACT

The data is a future life, science is a prove itself, data science is a prove future life itself, this paper present data science is a part of computer science engineers, they predict future life how to become successful. Data science used in several fields like medical, trades, digital marketing, stock marketing, education, government and private sectors many other fileds to determine how its successful to predict for socitial life. In data science use billions of technologies like machine learning, artificial intelligence, predictive modelling, cyber security, robotics, networks, industrial 4.0, big data analytics, iot and many more technologies. The data science is part of everyone's life how to face a challenges and solutions in this paper present in section 1, data science is a concept, section 2, applications and challenges, section 3, future technologies, section 4, data science case studies for engineers.

Keywords: Data Science, Technologies, Challenges, Application, Case Studies

I. INTRODUCTION: DATA SCIENCE IS A CONCEPT

Data science is a filed of study it involves extracting knowledge insights from noisy data into action data. The data science integratation between different sections, computer science, mathematics and business expertises. The three different types of data science analytics, based on the value and complex that is, descriptive, diagonstic, predictive and prescriptive analytics.



In descriptive analytics *(Mohsen Afsharian, 2024),* the data which is describes to what happend and make use of observaion and disribe the progress. In diagonstics analytics to understand the data and why took place for reason behind chosen the data. Predictive analytics to analyze and future based insight finally prescrptive to take care about before unavoidable action in future.

The Life Cycle of Data Sciecne

In data science mainly concentrate on collection of data basically different data set we have like raw data, images, audio, video, live streaming, sensor data etc. In (Fig 1) show the life cycle of data science firstly, collect the data

need to understand, clean the missing or noisy or unwanted data, explore pattern and classify data, apply the machine learning algorithm to visualize in different charts and function.

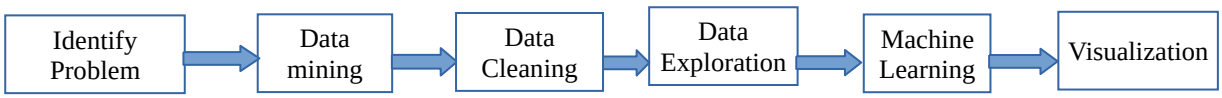


Fig 1: The data science life cycle.

II. RELATED WORK

The rapidly development of data science *(Zongben Xu, et.al, 2021)* in emerging technologies, understadning the patten of data is imporatant to improve the convergence of engineering management and data science technology. In this study the author discuss several significant challenges, interpretation and connotation of data science. The development and relationship of data science and direction of data science mainly focus on standard issues, security issues, technologies issues, storage and complex network issues. And future direction will gave suggestion to improve the data science in future development, science and technology.

The engineer design(ED) in data science to solve the problems and patterns within in requirement and particular constraint *(Filippo Chiarello, et.al, 2021). T*he author discuss some of the key factors in data science, they are, the engineer design process, problem and concept of data science design, enbodiment and detailed design, project and innovation development, design education, sustainability, tools of data science, applications, methods and data visulaization. The author discuss some the engineer design challeges are, the problem to develop data tools, top-down and bottom up approaches in data science, optimiation of genetic algorithm, CAD data source, NLP communication system to improve to serve the better ED process.

The data science is great research *(Inigo Nartinez, et.al, 2021)* for developing adavance analytics to improve data models and algorithms. Some of new technologies arise lack of chllenges to execute data science project in 1990's lack of technical issues, lack of clear vision objectives, low level ad-hoc projects etc. The author present new framework technology can be used for doing reserach as a roadmap for design a new data science methodologies.

The artificial intelligence and data science *(Muhammad Fakhrul Fafitra, et.al, 2024)* used in numerous field mainly focus on healtcare, education, finance, smart cities to find the traffic congestion that show promised AI application. The artificial intelligence and data science give more potential immensive business in a varity ways but need to safe improvement and implementation in ongoing projects. In digital world the data is wealth *(Iqbal H Sarker, 2021)*, in the current age of industrial revolution(Industry 4.0) useful insight from the data is decision making in smart application. The author said data science and advaned analytics made a good researcher, decision makers and application developer.

Data science is a emerging multi-disciplinary technology *(S Joe Qin, et.al, 2020)* for developing therotical and practical application. The author said the data science integrate with science and technologies, dynamic system and control modeling using artificial intelligence and machine learning technique.

III. APPLICATIONS OF DATA SCIENCE

Game Learning Analytics in data science

The game learning analytics *(Cristina Alona-Fernandez, et.al, 2019)* in data science using serious game designing can use algorithm or analysis technique, the stakeholder chose to benefits from the information with standard data sets.

Research and Development in Data Science

To increase research and development application in sustainable analysis *(Jennifer B. Dunn, et.al, 2021)* the great development of research, environemental and societal sustainability, multiple computer science engineering disciplines will bring power of data science very challenging. The data science used in industry 4.0 in future for design percepctive and improve the technology in *(Hélio Castro, et.al, 2022)* different sectors. In education, college and university the information science and data science are two different aspects to develop their own marketing for identifying opportnity in right way *(Chirag Shah, 2023).*

Health care and Medical in data science

In health care and medical field data science act as boon. Data science is used for detecting tumor, drug and medicine discovery, medical image analysis, genetic and genomics, virtual medical analysis and predicting diagosis etc.

Smart cities analytics and automated vehicle connectecd in data science

Now a days the automated vehicle are connected each other to communicate across the roads, the data science used for smart cities for road user to pridict real time traffic condition, accidents, traffic signal control system and any other issues are detecting for real time user.

Cyber security in data science

In advanced technology we need a qualified professional area of data science and cyber security to help hold digital information for securing our privacy data from attackers. So cyber security involves protection from vulnerabilities, identify and take action immedialty for protecting our data.

Logistics and Retail Management in data science

The various logistics and retail companies like FedEx, Amazon, Flipkart, etc. To make use data science helps to find the best route for deliver a product in a fixed time, the best easy mode of transporatation to reach destination.

IV. CHALLENGES

In real world the data science is challengable because the certain problem not suitable for developing project due to mismatch of quality of data. The different challenges we are facing in data quality and data quantity, security concern and privacy concern, lack of scalability and interdisciplinary data and problems with interpretability, impact problem for rapid development application and finding communication expandability.

V. TECHNOLOGIES USED FOR DATA SCIENCE

The data science is one of the most power tools for predict the future data analytics *(Table 1)*. The mainly use data science tools are Apache Spark, TensorFlow, Tableau, Jupyter Notebook, Excel, Python, Matplotlib, Scikit-learn, D3.js, MATLAB, PowerBI, Apache Hadoop, BigML, R and R Studio, KNIME, Numpy, Pandas, PyTorch, RapidMiner, Seaborn, SQL, Weka and SAS.

Tools	Overview	Key Features	Impact
Python	The Programing language for data sciecne due to simple, reliable and vast libraries for developing projects	Libraries like pandas, TensorFlow, Matplotlib, scikit learn for improved integration with data platform and performance optimization	Solve complex data analysis, Community driven support, Indispensable for data scientists.
R	Statistical Powerful BI and graphical capabilities	Package like tidyverse for data manipulation and ggplot2 for ploting	Benefits for academics and research development for statistical method
SQL	Standard query language for data base management system	Query integration with NoSQL Database	Education and industry based significant database. Handles large data efficiently
Apache Spark	Large scale data processing	Libraries like enhanced machine learning technique	Processing vast amount of data fastly used for real time

			applications
TensorFlow	Framework for Deep Learning	Focus on easy use of pre built models, API's	AI innovation drive with real world application
Tableau	Visual interface for complex data	Mobile support, Enhanced better integrated capabilities and real time data source	Used for educational and collaborative research industries.
Jupytern Notebook and Lab	Supportive web interface for data science	Collaborative tools and language support	Efficient project collaboration with change and managing work history
SAS	Powerful software suite tools for data analytics	Robust and Enhanced machine learning capabilities	Useful in healthcare and finantial data security
Microsoft Excel	Spreadsheet advanced tools	Integrated AI features for automation and Predictive tool	Fundamental tool for all fields due to data handling.

Table 1 : Data Science tools

VI. CASE STUDIES IN DATA SCIENCE

case 1: Predictive maintenance in manufacturing

the global industries use data science to implement predictive maintenance solutions because analying data from their industrial equipment, manufacturing product, industrial euipment, engines, turbins, engineer can predict the need for maintenance before a issues occurs. Heres how data science play a main role in manufacturing operations and predictive maintenance. The 30% reduction jet engines maintenance utilizing predictive analytics. The 15% increase in operational efficiency in data driven practices.

case 2 : Healthcare diagnostics

Data science enhance healthcare diagonstics by provinding personal treatment, the natural language processing using in IBM watson and path AI used for detecting cancer treatment and increase accuracy in 15% to 25%.

case 3: Smart cities and urban planning

In bangalore is the smart city using data science technique to optimize road planning for public service. The artificial intelligence and machine learning algorithm to detect traffic flow congestion and prediction. The data science use real time traffic management, reduce traffic in peak hours, consume fuel and time result in shorter.

case 4: Recommendation System

In all platform amazon, flipkart, ebay, the e-commerce have a raise to use data science to predict personalized shopping based on the recommendation system. Additionally, 40% and 22% due to fewer inquiries product recommendation and higher conversion in email make easier communication enhanced to user experiance.

case 5: Agriculture crop prediction

Farmers using data science to predict crop yields while conserving resources. The farmers have achieve 20% and 25% usage reduction in water and resulting cost saving and reduce environmental impact.

case 6: Fraud detection and Prevention system

The paypal and capital one is a leading online payments, the employs play a critical role in advanced to use data science technique for fraud tranction detection and prevention. The paypal and capital one use 75% reported their fraud prevention scheme measured for $2 billion losses due to unauthorized transaction in a year.

case 7: Energy or Power consumption prediction

EnergyOptiUS andcarbonSmart USA use data science to assist optimize energy consumption and control and monitor lighting system in real time. The energyoptiUS consume 20% power consumption in buildings and reduce carbon footprint. In carbonsmart USA use data science to reduce carbon footprint business, data driven sustainability and CO2 emission play a significant climate change role.

VII. CONCLUSION

Data science is not a buzz word, it's a extreme force that represent restructure the world and improve our daily lives. The real world case studies, challenges and applications as mentioned above incredible potential of data science in diverse domains from manufacturing, healthcare, smart cities and urban development, real time fraud detection and prevention, energy consumption and to agriculture beyond. In future direction, the technology advanced, as we can expect more innovative applications of data science will drive automatic progress in various sectors and solve some of the worlds most pressing challenges.

Declarations

conflict of interest: The authors declare no conflict of interest.

REFERENCES

[1] Zongben Xu, Niansheng Tang, Chen Xu, Xueqi Cheng, Data science: connotation, methods, technologies, and development, Data Science and Management 1 (2021) 32–37, https://doi.org/10.1016/j.dsm.2021.02.002

[2] Filippo Chiarello, Paola Belingheri, Gualtiero Fantoni, Data science for engineering design: State of the art and future directions, Computers in Industry 129 (2021) 103447, https://doi.org/10.1016/j.compind.2021.103447

[3] Inigo martinez, Elisabeth vilas, Igor G Olaizola, Data Science Methodologies: Current Challenges and Future Approaches, Big Data Research Volume 24, 15 may 2021, 10013, http://doi.org/10.1016/j.bdr.2020.10013

[4] Muhammad Fakhrul safitra, Muharman Lubis, Tien Fabrianti Kusumasari, Deyana Prastika Putri, Advancements in Artificial Intelligence and Data Science: Models, Applications and Challenges, Procedia Computer Science 234(2024) 381-388.

[5] Iqbal H Sarker, Data Science and Analytics: An overview from data driven smart computing, Decision making and Applications Perspective, SN Computer Science(2021) 2:377, https://doi.org/10.1007/s42979-021-00765-8

[6] S Joe Qin, Yining Dong, On data science for process systems modeling, control and operations, IFAC PaperOnLine 53-2(2020) 11325-11331.

[7] Cristina Alona-Fernandez, Antonio Calvo-morata, Manuel Freire, Ivan Martinez-ortiz, Baltasar Fernandez-Manjon, Applications of data science to game learning analytics data: A systematic literature review. Computers and Education, Volume 141, November 2019, 103612, https://doi.org/10.1016/j.compedu.2019.103612

[8] Jennifer B Donn, Prasana Balaprakash, Chaper 13 Research and development for increased application of data science in sustainability analysis, Data science applied to sustainability analysis, 2021, pages 282-292, https://doi.1016/B978-0-12-817976-5.00014-0

[9] Elizabeth Fernandes, Sergio Moro, Paulo Cortez, Data Science, Machine Learning and big data in digital journalism: Asurvey of state-of-art, challenges and opportunities. Experts Systems with Applications 221(2023) 119795, https://doi.org/10.1016/j.eswa.2023.119795

[10] Gilian Oliver, Data science and the information professions: challenges and opportuniities, Data and Information Management 7 (2023) 100030, https://doi.org/10.1016/j.dim.2023.100030

[11] Deepti chopra, Praveen arora, Swarm Intelligence in data science: Challenges, Opportunities and Applications, Procedia computer science 215(2022), 104-111, 4th International conference on innovative data communication technology and application.

[12] Dr K K Ramachandra, Data Science in 21st century: Evolution, challenges, and future directions, International journal of business data analytics, volume 1, issue 1, january-june 2024, pp 1-15, https://iaeme.com/Home/issue/IJBDA

[13] Inam Ullah, Deepak Adhikari, Xin Su, Francesco palmieri, Celimuge Wu, Chang Choi, Integration of data science with intelligent IOT(IIOT): current challenges and future perspectives, Digital communication and Networks(DCN), https://doi.org/10.1016/j.dcan.2024.02.007

[14] Hélio Castro, Filipe Costa, Luís Ferreira, Paulo Ávila, Goran D. Putnik, Manuela Cruz-Cunha, Data Science for Industry 4.0: A Literature Review on Open Design Approach, Procedia Computer Science 204 (2022) 877–884, https://doi.org/10.1016/j.procs.2022.08.106

[15] Chirag Shah, The past, the present, and the future of information and data sciences: A pragmatic view, Data and Information Management 7 (2023) 100028, https://doi.org/10.1016/j.dim.2023.100028

[16] Mohsen Afsharian, Data science essentials in business administration: A multidisciplinary perspective, Decision Analytics Journal 11 (2024) 100442, https://doi.org/10.1016/j.dajour.2024.100442