

Report for AI Survey

on

Data Science

by

Aarushi

24901301

MTech First Year

Centre for Artificial Intelligence

2024-26

**Submitted to
Dr. Satbir Singh**



**Dr. B R Ambedkar National Institute of Technology
Jalandhar – 144001 (Punjab)**

INDEX

S. No.	Topic	Page No.
1	Objective	3
2	Introduction of Data Science	3
3	Applications	3
4	Jobs and Skill Sets	4
5	Related Research Papers	7
6	Conclusion	12
7	References	13

OBJECTIVE

Create a report on a field of AI of choice, the current job openings and the skill set required for those jobs.

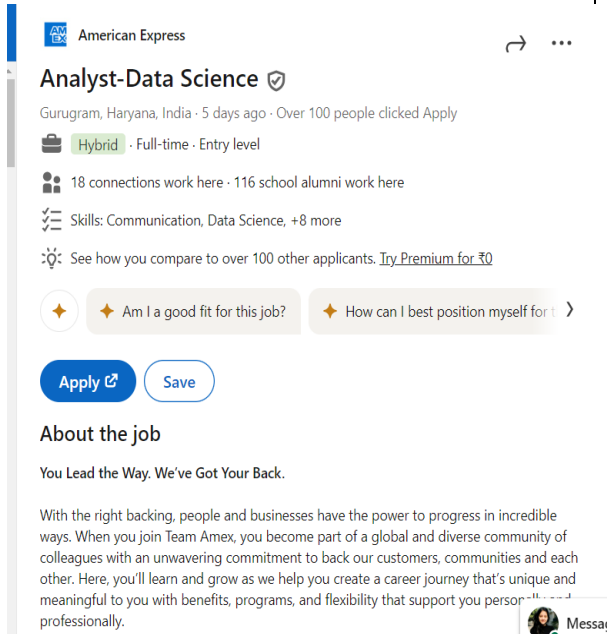
INTRODUCTION TO DATA SCIENCE






Data science is a multidisciplinary field that combines statistical analysis, machine learning, data visualization, and domain expertise to extract meaningful insights from structured and unstructured data. It leverages tools and techniques from mathematics, computer science, and programming to process, analyze, and interpret large datasets. The goal of data science is not just to understand data but also to make data-driven decisions and predictions that can solve real-world problems. By integrating methods such as data mining, predictive modeling, and big data technologies, data science plays a crucial role in industries like healthcare, finance, marketing, and technology. As organizations continue to generate vast amounts of data, the demand for skilled data scientists has grown significantly, making it a vital and evolving field in today's digital era.









APPLICATIONS OF DATA SCIENCE

1. Search Engines - Data Science helps deliver faster search results by analyzing website traffic to prioritize the most visited links. Example – When we search for something suppose “python courses ” then at that time on Internet we get the first link of a particular Course. This happens because that particular website is visited most in order to get information regarding python courses. So this analysis is done using Data Science, and we get the Topmost visited Web Links.
2. Transportation - Data Science enables driverless cars to reduce accidents by analyzing real-time data for safe driving decisions. Example, In Driverless Cars the training data is fed into the algorithm and with the help of Data Science techniques, the Data is analyzed like what is the speed limit in highways, Busy Streets, Narrow Roads, etc. And how to handle different situations while driving etc.
3. Finance - Data Science plays a crucial role in the stock market by analyzing past data to predict future stock prices. Data is analyzed in such a way that it makes it possible to predict future stock prices over a set timetable.
4. E-Commerce - E-commerce sites like Amazon, Flipkart, etc. uses data Science to make a better user experience with personalized recommendations. When we search for something on the E-commerce websites we get suggestions similar to choices according to our past data and also we get recommendations according to most buy the product, most rated, most searched, etc
5. Healthcare – Data Science plays a critical role in healthcare by enabling early detection of tumors through advanced image analysis techniques. It also supports the development of virtual medical bots that assist in patient care and diagnosis. These technologies enhance precision in medical treatments and improve overall patient outcomes.
6. Logistics - Various Logistics companies like DHL, FedEx, etc. make use of Data Science. Data Science helps these companies to find the best route for the Shipment of their Products, the best time suited for delivery, the best mode of transport to reach the destination, etc.

JOBS AND SKILL SETS

Website	Name of the Organization	Profile	Skill Set	Location
LinkedIn	<p>American Express[6]</p> 	Analyst-Data Science	<p>Minimum Qualifications</p> <ul style="list-style-type: none"> Master's Degree in quantitative field (e.g. Finance, Engineering, Physics, Mathematics, Computer Science or Economics) Strong programming skills Strong analytical/conceptual thinking acumen to solve unstructured and complex business problems Excellent written, oral communication skills. <p>Preferred Qualifications</p> <ul style="list-style-type: none"> Digital analytics experience and familiarity with paid marketing channels, digital marketing technology Experience with BIG DATA PROGRAMMING LANGUAGES (HIVE, PIG), PYTHON, JAVA 	Gurugram, Haryana, India
Indeed	S&P Global	Apprentice - Data Analyst	<ul style="list-style-type: none"> Course knowledge of Data Management, analytics, business intelligence, consulting, and/or strategy building. Knowledge in data visualization software, preferably Tableau . Knowledge in 	Gurugram, Haryana, India

	<p>Apprentice – Data Analyst</p> <p>S&P Global  · 3.9 ★</p> <p>Gurugram, Haryana</p> <p>Apply now  </p> <p>Job Company</p> <hr/> <p>Responsibilities :</p> <p>Work on visualization/reporting technologies, Scripting, and programming languages to produce one-time and cyclical deliverables for both internal and external clients.</p> <p>Collaborate in the software development lifecycle including Planning, Analysis, Design, Implementation, Testing, and Maintenance.</p> <p>Liaise with internal teams (e.g., developers and product managers) to identify system requirements.</p> <p>Identify process improvements within products and work to automate existing processes.</p> <p>Develop and apply testing processes for new and existing products or deliverables to meet customer needs.</p> <p>Anticipates product or deliverable release problems and takes corrective action, escalating as needed, to resolve and achieve commitments.</p>		<p>scripting language (e.g., Python) & SQL insights along with a solid understanding of relational databases.</p> <ul style="list-style-type: none"> • Problem-solving, self-motivation, and a drive for learning. • Ability to work with a diverse team. 	
Indeed	<p>Hewlett Packard</p> <p>Data Scientist</p> <p>Hewlett Packard  · 4.0 ★</p> <p>Bengaluru, Karnataka</p> <p>Location</p> <p> Bengaluru, Karnataka</p> <hr/> <p>Full job description</p> <p>Job Summary</p> <ul style="list-style-type: none"> • This role is responsible for conducting advanced analytics to generate breal insights enabling critical business decision making. The role works on buildi and improving medium complexity analytics models involving descriptive, d and predictive statistics. The role combines knowledge of data mining, stati modelling, mathematics, econometrics, advanced concepts and ability to ut data to generate actionable business insights. The role is involved in identifi setting up new data sources, and cleaning/organizing data. 	Data Scientist	<p>Preferred Certifications</p> <ul style="list-style-type: none"> • Programming Language/s Certification (SQL, Python, or similar) <p>Knowledge & Skills</p> <ul style="list-style-type: none"> • Algorithms • Artificial Intelligence • Big Data • Computer Science • Data Analysis • Data Mining • Data Science • Data Visualization • Deep Learning • Machine Learning • Natural Language Processing • Predictive Modeling • Python (Programming Language) • R (Programming Language) • SQL (Programming Language) • Statistical Modeling • Statistics • Tableau (Business Intelligence Software) <p>Cross-Org Skills</p> <ul style="list-style-type: none"> • Effective Communication 	Bengaluru , Karnataka

			<ul style="list-style-type: none"> Results Orientation Learning Agility Digital Fluency Customer Centricity 	
LinkedIn	<p>Spinny</p> <p> Product Analyst </p> <p>Spinny · Gurugram, Haryana, India (On-site)</p> <hr/> <p>About the job</p> <p>If you're as excited as we are about the power of data to transform businesses and build a superlative user experience, we have a role for you! We are a team of tinkerers, builders, and dreamers working toward the mission of making car ownership accessible, simple, and delightful. In this role, you will work closely with the product managers, marketing and engineering teams, business stakeholders to help build and improve products for our customers, employees, or partners using insights from data.</p> <p>Job location: Sector 48, Gurgaon Work from office: at least 3 days a week</p>	Product Analyst	<ul style="list-style-type: none"> Experience in data mining skills such as SQL, Python, and Excel Exposure to website analytics tools like Google Analytics or Adobe Analytics. Basic understanding of statistics such as hypothesis testing and Bayesian statistics. You should be able to generate insights and hypothesis from data and use them to drive business outcomes. The ability to work collaboratively, build relationships, and communicate effectively with stakeholders is crucial for this role. 	Gurugram, Haryana, India
LinkedIn	<p>Blackrock</p> <p> Applied Analytics , Analyst </p> <p>Mumbai Metropolitan Region · 2 weeks ago · Over 100 people clicked Apply</p> <p> Hybrid · Full-time · Entry level</p> <p> 14 connections work here · 1 company alum works here · 147 school alumni work here</p> <p> Skills: Data Analytics, Communication, +8 more</p> <p> See how you compare to over 100 other applicants. Try Premium for ₹0</p>	Applied Analytics, Analyst	<ul style="list-style-type: none"> Outstanding academic achievement: Undergraduate/Post Graduate Degree in Engineering, Statistics, Business Management or equivalent Strong hands-on experience in SQL and Power BI would be an added advantage. Strong inclination and progress towards learning and implementing newer technologies like Python. 	Mumbai, India

Paper1: [Data science and big data analytics: a systematic review of supply chain and logistics processes\[1\]](#)

Authors: Hamed Jahani, Richa Jain and Dmitry Ivanov

Source : Springer

Year: 2023

Introduction and Background

The paper addresses the extensive use of Data Science (DS) and Big Data Analytics (BDA) methodologies in Supply Chain and Logistics (SC&L). It highlights the fragmented nature of existing research and the need for a structured review to systematize DS&BDA applications in SC&L, focusing on efficiency, resilience, and sustainability

Key Findings

1. **Efficiency:** DS&BDA techniques significantly enhance operational efficiency by optimizing resource allocation, reducing lead times, and improving demand forecasting. Predictive analytics, such as machine learning algorithms, play a crucial role in these improvements¹.
2. **Resilience:** The study emphasizes the importance of DS&BDA in enhancing supply chain resilience. Techniques like real-time data analytics and predictive modeling help in identifying potential disruptions and developing mitigation strategies¹.
3. **Sustainability:** DS&BDA contribute to sustainability by enabling better resource management, reducing waste, and improving overall environmental impact. The paper discusses various case studies where DS&BDA have been successfully implemented to achieve sustainability goals.

Core Concepts in Data Science and Big Data Analytics

The authors begin by explaining the core concepts of data science and big data analytics, emphasizing the growing significance of these tools in industries. Data science is defined as the interdisciplinary field that extracts insights and knowledge from structured and unstructured data, while big data refers to high-volume, high-velocity, and high-variety datasets that traditional data processing tools cannot handle. Analytics methods, including predictive, prescriptive, and descriptive techniques, are highlighted as essential for enhancing supply chain operations. The integration of machine learning, artificial intelligence (AI), and data mining algorithms into logistics and supply chains has allowed businesses to predict demand patterns, optimize inventory levels, and mitigate risks.

Conclusion

The systematic review by Jahani, Jain, and Ivanov provides a comprehensive overview of how data science and big data analytics are transforming supply chain and logistics processes. From improving demand forecasting to enabling more sustainable practices, data-driven insights are helping businesses achieve operational excellence and gain a competitive edge. While there are challenges to overcome, the potential benefits of data science in supply chain management are immense, making it a critical area for future research and innovation.

Paper2: [Harnessing Big Data Analytics for Healthcare: A Comprehensive Review of Frameworks, Implications, Applications, and Impacts\[2\]](#)

Authors: Awais Ahmed, Rui Xi, Mengshu Hou, Syed Attique Shah, Sufian Hameed

Source : Springer

Year: 2023

Introduction

In the paper titled "**Harnessing Big Data Analytics for Healthcare: A Comprehensive Review of Frameworks, Implications, Applications, and Impacts**" by Awais Ahmed, Rui Xi, Mengshu Hou, Syed Attique Shah, and Sufian Hameed, the authors delve into the transformative capabilities of big data analytics within the healthcare sector. Published by Springer in 2023, this review article seeks to provide a thorough understanding of the frameworks that support big data analytics, the significant implications for healthcare practices, various real-world applications, and the broader impacts on the industry. By synthesizing current knowledge and advancements, the paper aims to offer a comprehensive guide for stakeholders in the healthcare ecosystem to effectively leverage big data for enhanced patient outcomes and operational efficiencies.

Key Frameworks

The authors highlight several critical frameworks that underpin big data analytics in healthcare. These frameworks include data acquisition, data storage, data management, data integration, and data analysis. Each framework is discussed in detail, outlining the technologies and methodologies used to handle the massive volumes of data generated in healthcare settings.

Applications

The applications section covers a wide range of uses for big data in healthcare. This includes predictive analytics for disease outbreaks, personalized medicine, electronic health records (EHRs), and patient monitoring systems. The authors provide examples of successful implementations and how these applications can lead to more efficient and effective healthcare delivery.

Conclusion

The paper concludes with a discussion on the future directions of big data analytics in healthcare. It calls for continued research and collaboration among stakeholders to address the challenges and fully realize the potential benefits of big data. The authors highlight the need for robust data governance frameworks, advanced analytical tools, and interdisciplinary approaches to harness the power of big data in healthcare effectively.

Paper3: [Data Science in the Design of Social Development Projects\[3\]](#)

Authors: Carlos Alberto González Almaguer, Veronica Saavedra Gastelum, Arturo Gonzalez de Cosio, Lourdes Jazmin Muciño Garcia, Claudia Zubieta Ramirez

Source : IEEE

Year: 2023

Introduction

In the paper titled "**Data Science in the Design of Social Development Projects**" by Carlos Alberto González Almaguer, Veronica Saavedra Gastelum, Arturo Gonzalez de Cosio, Lourdes Jazmin Muciño Garcia, and Claudia Zubieta Ramirez, published by IEEE in 2023, the authors explore the application of data science methodologies in the design and implementation of social development projects. The paper aims to highlight the potential of data science to address social issues and improve the effectiveness of development initiatives.

Key Frameworks

The authors outline several frameworks for integrating data science into social development projects. These include data collection and analysis, predictive modeling, and impact evaluation. The paper highlights the need for robust data management practices and the use of advanced analytics to inform decision-making.

Applications

The applications section showcases various projects where data science has been successfully applied to social development. Examples include using predictive analytics to identify at-risk populations, designing targeted interventions to improve health outcomes, and evaluating the impact of social programs to inform policy decisions.

Conclusion

The paper concludes with a call for increased collaboration between data scientists, social development practitioners, and policymakers. It emphasizes the need for interdisciplinary approaches to harness the full potential of data science in addressing social challenges. The authors highlight the importance of building capacity and infrastructure to support data-driven social development projects.

Overall, the paper provides a comprehensive overview of how data science can be applied to social development projects, the challenges faced, and the transformative impacts it can have on society.

Paper4: [Environmental Resilience Technology: Sustainable Solutions Using Value-Added Analytics in a Changing World \[4\]](#)

Authors: E. Natasha Stavros, Caroline Gezon, Lise St. Denis, Virginia Iglesias, Christina Zapata, Michael Byrne, Laurel Cooper, Maxwell Cook, Ethan Doyle, Jilmarie Stephens, Mario Tapia, Ty Tuff, Evan Thomas, S. J. Maxted, Rana Sen, and Jennifer K. Balch

Source : Applied Sciences, Volume 13

Year: 2023

Introduction

In the paper titled "**Environmental Resilience Technology: Sustainable Solutions Using Value-Added Analytics in a Changing World**" the authors explore the role of data-driven technologies in enhancing environmental resilience. The paper discusses how value-added analytics can be used to develop sustainable solutions to address the challenges posed by global climate change and environmental extremes.

Key Frameworks

The authors outline several frameworks for implementing environmental resilience technologies, including data acquisition, data integration, and value-added analytics. These frameworks are designed to enable society to become more resilient to changing environmental conditions.

Applications

The applications section showcases various case studies where data science has been successfully applied to environmental resilience projects. Examples include wildfire response, risk mitigation, and recovery efforts. The authors provide insights into how these applications can lead to more effective and sustainable environmental management.

Conclusion

The paper concludes with a discussion on the future directions of environmental resilience technologies. It calls for continued research and collaboration among stakeholders to address the challenges and fully realize the potential benefits of data-driven solutions. The authors highlight the importance of interdisciplinary approaches and the need for advanced analytical tools to support environmental resilience efforts.

Overall, the paper provides a comprehensive overview of how data science can be applied to environmental resilience projects, the challenges faced, and the transformative impacts it can have on society.

Paper5: [Climate Change Mitigation and Adaptation using AI: Developing Sustainable Solutions\[5\]](#)

Authors: Moses Alabi, Vincent Gerald

Source : Research Gate

Year: 2024

Introduction

In the paper titled "**Climate Change Mitigation and Adaptation using AI: Developing Sustainable Solutions**" by Moses Alabi and Vincent Gerald, published on Research Gate in 2024, the authors explore the potential of artificial intelligence (AI) in addressing climate change. The paper discusses how AI can be leveraged to develop innovative and sustainable solutions for mitigating and adapting to the impacts of climate change.

Key Frameworks

The authors outline several frameworks for implementing AI in climate change initiatives, including data collection, data analysis, predictive modeling, and decision support systems. These frameworks are designed to enhance the efficiency and effectiveness of climate change mitigation and adaptation efforts.

Applications

The applications section showcases various case studies where AI has been successfully applied to climate change mitigation and adaptation. Examples include using AI for weather forecasting, optimizing renewable energy systems, and developing early warning systems for natural disasters. The authors provide insights into how these applications can lead to more resilient and sustainable communities.

Impacts

Finally, the paper discusses the broader impacts of AI on climate change mitigation and adaptation. These impacts include improved disaster response, enhanced decision-making processes, and the ability to adapt to environmental changes. The authors argue that AI has the potential to transform climate change efforts by providing actionable insights and fostering innovation.

Conclusion

The paper concludes with a discussion on the future directions of AI in climate change mitigation and adaptation. It calls for continued research and collaboration among stakeholders to address the challenges and fully realize the potential benefits of AI-driven solutions. The authors highlight the importance of interdisciplinary approaches and the need for advanced analytical tools to support climate change initiatives. Overall, the paper provides a comprehensive overview of how AI can be applied to climate change mitigation and adaptation, the challenges faced, and the transformative impacts it can have on society.

CONCLUSION

This report has highlighted the multifaceted role of data science and artificial intelligence across various domains, emphasizing their transformative impact on industries, societal challenges, and environmental issues. From enhancing operational efficiency in supply chains to advancing personalized healthcare and addressing climate change, the integration of advanced analytics and AI technologies is proving essential in solving complex real-world problems. The analysis of job trends and skill requirements underscores the growing demand for professionals skilled in data science and AI, particularly in areas such as predictive modeling, big data analytics, and programming languages like Python, SQL, and R. Companies are seeking candidates with strong analytical abilities, technical expertise, and collaborative skills to drive innovation and deliver data-driven solutions. Insights from reviewed research papers further reinforce the potential of data science and AI to address critical challenges. In supply chains, they optimize resources and resilience; in healthcare, they enhance patient outcomes; in environmental efforts, they foster sustainability and resilience; and in social development, they improve the effectiveness of interventions.

Looking ahead, the continued integration of data science and AI in diverse sectors will require interdisciplinary collaboration, robust data governance, and ongoing skill development to maximize their benefits. The transformative capabilities of these technologies position them as cornerstones of innovation in a rapidly evolving digital landscape, paving the way for a more informed, sustainable, and resilient future.

REFERENCES

- [1] Jahani, H., Jain, R. & Ivanov, D. Data science and big data analytics: a systematic review of methodologies used in the supply chain and logistics research. *Ann Oper Res* (2023). <https://doi.org/10.1007/s10479-023-05390-7>
- [2] A. Ahmed, R. Xi, M. Hou, S. A. Shah and S. Hameed, "Harnessing Big Data Analytics for Healthcare: A Comprehensive Review of Frameworks, Implications, Applications, and Impacts," in *IEEE Access*, vol. 11, pp. 112891-112928, 2023, doi: 10.1109/ACCESS.2023.3323574.
- [3] C. A. G. Almaguer, V. S. Gastelum, A. G. d. Cosio, L. J. M. Garcia and C. Z. Ramirez, "Data Science in the Design of Social Development Projects.," *2023 Future of Educational Innovation-Workshop Series Data in Action*, Monterrey, Mexico, 2023, pp. 1-5, doi: 10.1109/IEEECONF56852.2023.10104990.
- [4] Stavros, E.N.; Gezon, C.; St. Denis, L.; Iglesias, V.; Zapata, C.; Byrne, M.; Cooper, L.; Cook, M.; Doyle, E.; Stephens, J.; et al. Environmental Resilience Technology: Sustainable Solutions Using Value-Added Analytics in a Changing World. *Appl. Sci.* 2023, *13*, 11034. <https://doi.org/10.3390/app131911034>
- [5] Alabi, Moses & Gerald, Vincent. (2024). Climate Change Mitigation and Adaptation using AI: Developing Sustainable Solutions.
- [6] Job Source:
<https://www.linkedin.com/jobs/view/4046849998/?alternateChannel=search&refId=adfe2171-c6b0-4ae1-8011-a348b0d7461e&trackingId=Lab5fQTqTj6ffJbxPLO5BQ%3D%3D>