

## CASE STUDIES AND EDUCATION IN DATA SCIENCE IN HIGHER EDUCATION INSTITUTIONS

**Dogiparthi Shravan Kumar**

Lecturer in Computer Science

Gov City College(A),

Hyderabad.

### Abstract:

While the Data Science industry is booming, it is still very nascent but very promising. This industry is expected to grow by 26% by 2026, The most common source available now is our GPRS services. So, when we activate the GPRS for our road travel, it uses various branches of data science to seek out the shortest route, estimate time of arrival, nearby places for petrol re-fill and eating outlets. All these activities are assisted by the data science software running in the background. To summarise, the Data Science team everywhere helps the organizations manage, tackle, analyze their data every day. To simplify, Data science is a field that helps to structure the data using the Data Mining Technique.

### Introduction

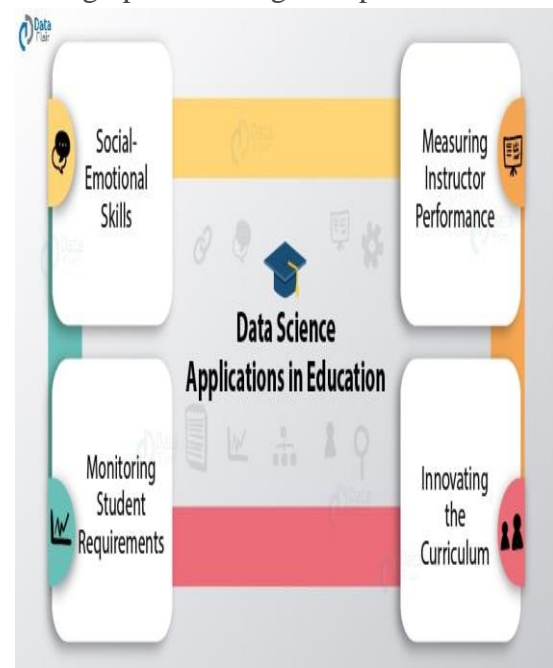
Data Science has spread its branches through several quintessential fields of the world today. It has emerged out as a global phenomenon that has revolutionized industries and has increased their performances substantially. As a matter of fact, Data Science has extended itself to incorporate education under its wing

### Data Science in Education

Education is the key to shaping the lives of people. It holds the power to transform and enrich the lives of people. Since the dawn of civilization, humans have evolved through education and have developed mechanisms to improve education.

In the 21st century, where data is omnipresent in every walk of life, education is no exception. With

advancements in computing techniques, it is possible to imbibe all the information through powerful big-data platforms.



### Applications of Data Science in Education

#### 1. Social-Emotional Skills

Social-Emotional Skill is an important area that needs to be developed through education. Through this, a child learns to acquire a capacity to understand, analyze, express and manage emotions. He also learns how to develop a relationship with others.

Facilitating growth in social-emotional skills is an important task of educational institutes. This is an example of a non-

academic skill that plays a major role in defining the learning capabilities of the students.

Previously, there were various statistical surveys that would assess these social-emotional skills. However, with the advancements in computational methodologies, it is possible to gather a large amount of data. With the formalized knowledge discovery models in Data Science and Data Mining techniques, it is possible to gather such large information and incorporate it with the existing tools to produce better results.

Furthermore, with the data that is extracted, data scientists can apply various predictive analytical techniques to help the teachers in understanding the motivation of the students for studying the course.

## 2. Monitoring Student Requirements

There are several evaluation and assessment techniques that are utilized by educational institutes. However, such traditional methodologies were often unable to capture and encapsulate all the important trends and patterns of student services.

Furthermore, most of the assessment techniques were not in real time. With the advancements in Big Data analytics, it is now possible for the teachers to scrutinize their student requirements based on their performance and reviews. As a result of monitoring student requirements, teachers are able to provide appropriate responses and even change their teaching methodologies to meet student expectations. Many times, teachers have an unconscious bias towards certain students.

A data platform will treat its users with a zero bias, meaning that there will not be any bias in the evaluation of student performance. This will provide an equal platform for all the students to engage and develop their skills.

## 3. Innovating the Curriculum

Various Universities have to keep themselves updated with the demands of the industry so as to provide appropriate courses to their students. Furthermore, it is a challenge for the universities to keep up with the growth of industries. In order to accommodate this, Universities are using Data Science systems to analyze growing trends in the market. Using various statistical measures and monitoring techniques, data science can be useful for analyzing the industrial patterns and help the course creators to imbibe useful topics. Furthermore, using **predictive analytics**, universities can analyze demands for new skill sets and curate courses that address them.

## 4. Measuring Instructor Performance

The performance of students depends on the teachers. While there are many assessment techniques that have been used to assess the performance of teachers, it has been mostly manual in nature.

**For example**, student reviews about the performance of teachers make, have been the standard tool for quantifying teaching methodologies. However, all these techniques are not an efficient method and usually take time to assess. Furthermore, reading student reviews and creating an analogy is a tiresome task. With the breakthrough in data science, it is possible to keep track of the teacher performance. This is not only valid for recorded data but

also real-time data. As a result, with real-time monitoring of teachers, rigorous data collection is possible, along with its analysis.

Furthermore, we can store and manage unstructured data like student reviews on a big data platform. And, with Natural Language Processing, it is possible to analyze the sentiments of the reviews and provide a comprehensive analysis of teacher performance.

### **1 Improve learning- Data Science in Education**

Yes, what can be other than learning improvement in the use case of data science in the education industry. We can use multiple machine learning models and statistical techniques to help students improve the way they learn.

With the combination of big data analytics and **data science**, we can gather the data and process it to find many important factors. Some of those can be-

- What are the areas students are lacking
- What are their pain points like is it concepts, labs, or more
- How timing can impact the adaptive learning
- Maintaining the better environment in the premise
- Improving the course content
- Understanding the impact of professors on the students and more

### **2 Career Path Prediction**

- Machine learning can be used to predict the career path for the students. Along with it, it can also help in selecting the best colleges depending on the various factors.

### **3. Personalized Classroom**

Although a lot has been done in the classroom personalization with the help of data science and machine learning. But a lot needs to be done further. And so, we have added classroom personalization as one of the use cases of data science in education.

### **4 Connecting the company with right set of influencers**

Data science in education can be further generalized for the educational product company who want to reach a wider set of audiences. For example, let's say there is a company who wishes to reach a certain set of audience as a possible customer. For this, they might want to connect to some good influencers. Those can help the company reach the wider and potential set of customers.

There are multiple factors which identify the suitable influencers. For example, if the organization wants to reach out to the Instagram influencers, they might need to check multiple factors like number of followers, reach, engagement, niche and more. You can check more about such influencers in the fashion domain here.

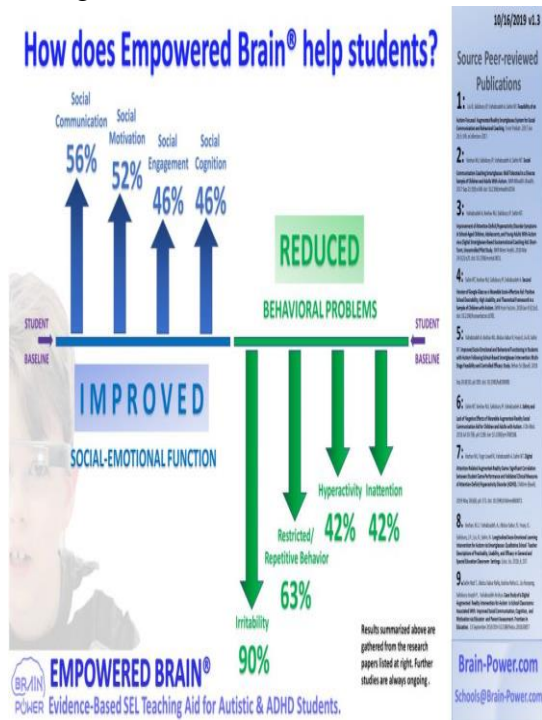
Although just having more followers doesn't guarantee for the best in the domain. There can be spammed followers as well. And here comes the role of machine learning algorithms which can detect such spam profiles. This way, the company can refrain from wasting the budget and time on such influencers.

### **5 Improve Grading System**

Issue in the grading system is always a questionable topic. Many education professionals have raised questions on a biased grading system. You might have found many scenarios when for the same

student, teacher-1 has given A grade while another teacher might have given grade B.

Such kind of biasing happens as every teacher has different ways of grading and is the problem. This is the same problem which occurs when you will see one of the school's students is getting a higher grade while other schools' are a little lower. This is because of the nature of teachers who have graded the students.



analysis, processing, sharing, and storage. Cloud computing services are generally easy for people and organizations to use, they are accessible over the Internet through a variety of platforms (workstations, laptops, tablets, and smart phones), and they may be able to accommodate spikes in demand much more readily and efficiently than in-house computing services.

**Advantages:**

1. Educational data science would help and train the ‘educators’ or ‘teachers’ where they will be able to enhance their teaching style and understand various techniques that interest the students more.
2. Educational data science will encourage educators to include data visualization, data reduction and description, and prediction tasks.
3. Data reduction will simplify the process of grading and assignments for the students.
4. The data visualization process will help students attract complex data in a more straightforward form and be taught in a storytelling format.

Cloud computing services are application and infrastructure resources that users access via the Internet. These services, contractually provided by companies such as Apple, Google, Microsoft, and Amazon, enable customers to leverage powerful computing resources that would otherwise be beyond their means to purchase and support. Cloud services provide services, platforms, and infrastructure to support a wide range of business activities. These services support, among other things, communication; collaboration; project management; scheduling; and data

The schools and other online programme providers must understand the contextualization of education: what works for whom, when and where – this will keep motivating the programme’s instructor to find new ways of teaching.

**Data Science in Ecommerce**

In the e-commerce sector, big data analytics can assist in customer analysis, reduce operational costs, forecast trends for better sales, provide personalized



shopping experiences to customers, and many more.

Amazon uses data science to personalize shopping experiences and improve customer satisfaction. Amazon is a globally leading eCommerce platform that offers a wide range of online shopping services. Due to this, Amazon generates a massive amount of data that can be leveraged to understand consumer behavior and generate insights on competitors' strategies. Amazon uses its data to provide recommendations to its users on different products and services. With this approach, Amazon is able to persuade its consumers into buying and making additional sales. This approach works well for Amazon as it earns 35% of the revenue yearly with this technique. Additionally, Amazon collects consumer data for faster order tracking and better deliveries.

Similarly, Amazon's virtual assistant, Alexa, can converse in different languages; uses speakers and a camera to interact with the users. Amazon utilizes the audio commands from users to improve Alexa and deliver a better user experience

### **BioTech**

How AstraZeneca harnesses data for innovation in medicine

AstraZeneca is a globally known biotech company that leverages data using AI technology to discover and deliver newer effective medicines faster. Within their R&D teams, they are using AI to decode the big data to understand better diseases like cancer, respiratory disease, and heart, kidney, and metabolic diseases to be effectively treated. Using data science, they can identify new targets for

innovative medications. In 2021, they selected the first two AI-generated drug targets collaborating with BenevolentAI in Chronic Kidney Disease and Idiopathic Pulmonary Fibrosis.

Data science is also helping AstraZeneca redesign better clinical trials, achieve personalized medication strategies, and innovate the process of developing new medicines. Their Center for Genomics Research uses data science and AI to analyze around two million genomes by 2026. Apart from this, they are training their AI systems to check these images for disease and biomarkers for effective medicines for imaging purposes. This approach helps them analyze samples accurately and more effortlessly. Moreover, it can cut the analysis time by around 30%.

AstraZeneca also utilizes AI and machine learning to optimize the process at different stages and minimize the overall time for the clinical trials by analyzing the clinical trial data. Summing up, they use data science to design smarter clinical trials, develop innovative medicines, improve drug development and patient care strategies, and many more.

### **Pharmaceutical**

Driving innovation with NLP: Novo Nordisk

Novo Nordisk uses the Linguamatics NLP platform from internal and external data sources for text mining purposes that include scientific abstracts, patents, grants, news, tech transfer offices from universities worldwide, and more. These NLP queries run across sources for the key therapeutic areas of interest to the Novo Nordisk R&D community. Several NLP algorithms have been developed for

the topics of safety, efficacy, randomized controlled trials, patient populations, dosing, and devices. Novo Nordisk employs a data pipeline to capitalize the tools' success on real-world data and uses interactive dashboards and cloud services to visualize this standardized structured information from the queries for exploring commercial effectiveness, market situations, potential, and gaps in the product documentation. Through data science, they are able to automate the process of generating insights, save time and provide better insights for evidence-based decision making.

### **Data Science in Entertainment Industry**

Due to the Pandemic, demand for OTT (Over-the-top) media platforms has grown significantly. People prefer watching movies and web series or listening to the music of their choice at leisure in the convenience of their homes. This sudden growth in demand has given rise to stiff competition. Every platform now uses data analytics in different capacities to provide better-personalized recommendations to its subscribers and improve user experience.

### **What Are the Skills Required for Data Scientists?**

Data scientists play an important role in the data science process as they are the ones who work on the data end to end. To be able to work on a data science case study, there are several skills required for data scientists like a good grasp of the fundamentals of data science, deep knowledge of statistics, excellent programming skills in Python or R, exposure to data manipulation and data analysis, ability to generate creative and

compelling data visualizations, good knowledge of big data, machine learning and deep learning concepts for model building & deployment. Apart from these technical skills, data scientists also need to be good storytellers and should have an analytical mind with strong communication skills.

### **FUTURE OF DATA SCIENCE AND AI IN INDIA**



Artificial Intelligence (AI) is the potential of machines to impersonate the capabilities of the human mind. These machines are enabling high-level cognitive processes like thinking, perceiving, learning, problem-solving and decision-making with advancements in data collection, aggregation, analytics, and computer processing power. The goal of AI systems is to tackle complex problems in similar ways to human logic and reasoning. The market for AI is expected to reach US\$ 7.8 billion by 2025 in India. According to the State of the Education report (SOER) 2022, AI will grow at a rate of 20.2% compound annual growth rate (CAGR). In recent years, all nations place the highest priority on strengthening educational standards and student learning results. By 2030, AI in education systems will contribute considerable efforts to achieve sustainable development goals in India along with help in addressing issues

related to equality, equity, and inclusion in education. AI will complement and supplement human intelligence and enrich the way people live and work.

AI has been the most revolutionary creation and is expected to have a prominent impact on the evolution of mankind. The global AI market in 2021 was nearly US\$ 59.67 billion and it is projected to grow at a CAGR of 39.4% to reach US\$ 422.37 billion by 2028. While the AI market in India is projected to grow at a CAGR of 20.2% to reach US\$ 7.8 billion by 2025 from US\$ 3.1 billion in 2020. The number of AI start-ups increased 14 times from 2000 till September 2022. Artificial intelligence has been gaining massive traction due to the enormous change in business operations and fast-paced technological advancement. The trend will only grow further in the coming years.

Experts have predicted that, by 2030, there will be a 31.4% increase in jobs related to data science and mathematical science, mostly AI-based.

Data science has been one of the most extensive applications of artificial intelligence as AI can provide massive analytical power. It facilitates extracting insights and patterns from a large data set and using it to make predictions on possible outcomes. Some common systems used for data analysis involve Google Analytics, automation platforms, business intelligence systems, content management systems, and CRMs. Data science users often use AI to get more refined value from already existing data and make valuable predictions.

Data science and AI are currently reigning technologies which have conquered industries around the world owing to the

massive explosion in data and the increasing need for businesses to rely on data for major decision-making. Data scientists help businesses analyse large amounts of data from numerous sources, assess their performance, and recommend necessary adjustments to improve it. Additionally, they support the product development team in customising goods and services by examining consumer behaviour. The extensive analysis of data through AI helps to save a lot of money and time for businesses.

The emerging technology has made it all-pervasive. It is not only for big businesses and industries instead it impacts our daily lives as well. We come across the influence of these technologies throughout the day, from weather forecasts, filtering spam emails, enabling search predictions in Google, personalizing social media feeds, and voice recognition such as Apple's Siri, Google Now, Microsoft's Cortana or Amazon's Alexa. The data science technologies also help OTT and other video streaming platforms like Netflix and Spotify to extract customer preferences and their viewing patterns to develop and curate highly targeted shows. Various e-commerce giants such as amazon, Flipkart, Meesho, etc., use customer data such as buying patterns, tastes, and preferences to present customised shopping recommendations.

### **Future of Artificial Intelligence (AI) in India**

AI is one of the emerging industries that is turning out to be a proxy for human brains. It performs various business functions without a human intervention like customer interaction, creating brand awareness on social media, etc. AI is

widely transforming various sectors such as healthcare, insurance, finance, marketing, etc., by automating their processes. It helps these sectors to analyse records, conduct market research, interact with potential customers, etc. AI has huge potential across the globe. India is the fastest-growing economy with the second-largest population in the world and has a significant stake in the AI revolution.

Considering the potential of AI to transform the economy, the finance minister of India in budget 2018-19 mandated NITI Aayog to establish a national program on AI. It was organised with a view to guiding the research and development segment about new and emerging technologies.

In 2022-23, the public funding for the digital India mission increased by 67% to reach US\$ 1.29 billion (Rs. 10,676 crore). This mission involves a plan for the effective utilization of AI to promote financial inclusion, supplement the education sector, and transform the urban infrastructure. States such as Tamil Nadu, Punjab, Uttar Pradesh, and Telangana are already utilising AI-based tools to support law and order, increase agricultural productivity, and improve healthcare delivery.

AI helps to contribute to various sectors such as agriculture and healthcare. India had an estimated 1,000 agriculture start-ups working with the government as of March 2022 in the agrotech segment.

The National Agriculture Market and eNAM, an electronic trading platform across India, are two of the top government assistance programmes that have helped to improve the agricultural system. Using sustainable technologies, the National Sustainable Agriculture

Mission seeks to increase agricultural output. A national e-Governance Plan for agriculture gives funding for cutting-edge technology like blockchain, machine learning, drones, and AI top priority. Farmers in this area get the benefits of using GPS, GIS, and satellite imagery. GPS-enabled devices can be used to guide drones and help in monitoring and implementing better irrigation practices.

Farmers who deal with periodic yield monitoring and inconsistency can benefit from start-ups that focus on GPS, GIS, and satellite images. Data about crop health such as the type and extent of disease manifestation can also be recorded to improvise crop quality. These data are helpful in supporting decisions about irrigation and fertilizer requirements. As a result, farmers can take considerable action to mitigate damage and associated costs. The Indian agriculture sector accounts for around 19% of the country's greenhouse gas emissions. Emerging technologies have helped in the control of various imprudent and polluting practices. Start-ups such as CropIN, AgroStar, DeHaat, Fasal, and SatSure are addressing the issues.

The Central Board of Secondary Education (CBSE) in accordance with National Education Policy (NEP) has introduced artificial intelligence as a subject in class IX and class XI in their affiliated schools implementing from the academic year 2020-21. During a discussion in the Lok Sabha in August 2022, Minister of State for Education Annpurna Devi highlighted AI initiatives like the Diksha portal. This portal uses artificial intelligence methods to offer self-paced learning, and it is designed using open-source software to offer content for school education in states



as well as UTs. It also provides QR-coded energised texts for all grades (one nation, one digital platform). As of March 28 2023, DIKSHA has 16.82 million students enrolled in various courses. A total of 9.32 million learning sessions have been attended in Karnataka followed by Rajasthan with 6.10 million and Odisha accounts for 4.45 million total learning sessions. The idea of cloud-based education in India is expanding widely as it helps to improve physical and digital access to resources. Companies like Miko, an artificial intelligence (AI) driven companion robot for kids, offer services including chatting, reacting, instructing, amusing, and comprehending the child's needs, emotions, and likes and dislikes. These businesses, along with other start-ups, gained traction in the Covid, creating an online ecosystem enabling kids to learn more quickly. There are many AI start-ups developed in 2022 which contribute to the nation's education segment such as HackerRank, iNurture Square Panda etc. Healthcare is also one of the sectors which have included AI to improve performance across the sector. One of the recent examples is the collaboration of Google with Apollo Hospitals in India to improve the deep learning models in x-rays and other diagnostic purposes. The issue of poor availability of better healthcare facilities in rural areas is being analysed by AI. The early and rapid detection of these issues can be a powerful tool for targeted public health interventions, particularly in rural areas. The adoption of enhanced technologies and automated intervention provides opportunities to bridge existing gaps in the healthcare sector. Companies like Google, Microsoft, Meta, and Apple have spent around US\$ 3 billion in 2021 in

the healthcare sector to amplify the growth with start-ups such as Pharmeasy, HealthifyMe, Healthplix, DocTalk, etc. The algorithms of AI have augmented the healthcare space ranging from early disease diagnosis, drug recovery trials, and precision in patient monitoring to self-care. During complete lockdown in India, AI-powered start-up MyHealthcare which is recognised by NASSCOM built AI-based solutions by adopting speciality care EMR solutions, voice-based CPOE and AI-enabled CDSS that helps to deliver personalised healthcare services to their patients. There has been the integration of AI with diagnostic algorithms for screening diseases ranging from cancer, and diabetic retinopathy to cardiovascular disease.

The Indian defence industry is working towards transforming the armed forces into one of the most advanced in the world. The adoption of various technologies based on AI will revolutionise the Indian Military and help India to become one of the biggest defence product markets. This collaborative effort among the public and private sectors of industry, research organisations, academic institutions, start-ups, and innovators has contributed to the development of numerous innovative technological products based on AI in the fields of data, logistics, surveillance, weapons, and many others. The introduction of autonomy in weapon systems, in Intelligence, Surveillance and Reconnaissance (ISR) data management, can be a huge asset in stopping terrorism, installing counter-terrorism measures, and protecting soldiers. In fact, AI in defence can change combat and conflict at the deepest levels.

### Conclusion of Data Science in Education

This was all about use cases of data science in the education system which help the education system improve the performance. Apart from the above top 5 data science applications in education there can be many more use cases which you can choose and implement.

If you have implemented any data science application in the education system, please share with us in the comment

### References

- *Klašnja-Milićević, A., Ivanović, M., & Budimac, Z. (2017). Data science in education: Big data and learning analytics. Computer Applications in Engineering Education, 25, 1066–1078. <https://doi.org/10.1002/cae.21844>*
- *López-Zambrano, J., Lara, J. A., & Romero, C. (2021). Improving the portability of predicting students' performance models by using ontologies. Journal of Computing in Higher Education. <https://doi.org/10.1007/s12528-021-09273-3>*
- *Maatuk, A. M., Elberkawi, E. K., Aljawarneh, S., Rashaideh, H., & Alharbi, H. (2021). The COVID-19 pandemic and E-learning: Challenges and opportunities from the perspective of students and instructors. Journal of Computing in Higher Education. <https://doi.org/10.1007/s12528-021-09274-2>*
- *Mitrofanova, Y. S., Sherstobitova, A. A., & Filippova, O. A. (2019). Modeling smart learning processes based on educational data mining tools. In V. Uskov, R. Howlett, & L. Jain (Eds.), Smart Education and e-Learning 2019. Smart Innovation, Systems and Technologies. (Vol. 144). Springer. [https://doi.org/10.1007/978-981-13-8260-4\\_49](https://doi.org/10.1007/978-981-13-8260-4_49)*
- *Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. Wires Data Mining and Knowledge Discovery, 10, e1355. <https://doi.org/10.1002/widm.1355>*