PLAY AND MAKE MONEY FROM THE OCEAN OF DATA

The hiring market for data scientists and analytics professionals has gone into an overdrive and institutions are now offering new courses in Data Science...

by Abhay Anand

Big Data, Artificial intelligence, Machine Learning, Data Science and Data Analytics are some of the words doing rounds in every field these days. Most industry reports point towards huge demand for professionals with an understanding of Data Science. A LinkedIn survey has put Machine learning (ML) Engineers and Data Scientists among the 'Emerging Jobs' list. Data Science has emerged as one of the most exciting careers with indications that the current number of machine learning engineers is around 10 times that of five years ago.

What is Data Science

Data Science requires accumulation of data, arranged and analysed to examine its effect on businesses. Google describes a Data Scientist as "a person employed to analyze and interpret complex digital data, such as the usage statistics of a website, especially in order to assist a business in its decision-making". She/he chooses and builds appropriate algorithms and models to analyze the data in a better manner and uncover insights from it.

Increasingly, higher education institutions have also realised the importance of the career opportunities in this field and have either tweaked their existing programmes or have started a full-time course in Data Science or Data Analytics.

Leading the pack are the Indian Institutes of Technology (IITs), which have started B.Tech course in Data Analytics. IIT Madras has started an interdisciplinary dual degree program in data science. IIT Hyderabad has kicked off an M.Tech in Data Science, IIT Kanpur and IIT Mandi are also planning to launch courses in this area.

Data Science versus traditional B.Tech CS/IT

The traditional Computer Science (CS) course is the study of computers design, architecture and its application. It includes hardware, software, networking and internet. Information technology (IT) deals with the application of computing technology to real-life processes. IT is just the selection of the proper software or hardware for a given task.

On the other hand, a dedicated course in Data Analytics/Science programme is the study of various types of data; structured, semi-structured and unstructured data in any form or formats available in order to get some information out of it. The course consists of different technologies used for data mining, data storing, data purging, data archival, data transformation etc.

Dr. Vineeth N. Balasubramanian, Associate Professor, Department of Computer Science and Engineering, IIT Hyderabad, explains, "It is envisaged that students graduating from these programs will apply their training to become leaders in the industry and academia and contribute to building a better future for India and the world. The basic aim is to create a complete ecosystem for Artificial Intelligence Academics and Research at IIT Hyderabad. This involves B.Tech., M.Tech. and different Minor Programs in AI. Moreover, the R&D will be strongly entwined with academics."

The program at IIT Hyderabad is also unique as it aims to provide a holistic view to students. It comprises algorithms from the Computer Science Department, Signal Processing from Electrical Engineering Department, Robotics from the Mechanical Engineering Department and Mathematical Foundations. The course will also focus on application verticals such as healthcare, agriculture, smart mobility, among many others.

Further explaining the difference

Prof. Balaraman Ravindran,
Head, Robert Bosch Centre for Data Science and Artificial Intelligence
IIT Madras

The program we have launched is an interdisciplinary dual degree program in data science. The goal of this program is to train people with competency in a core discipline to become better data scientists

Shutterstock

between the traditional course of CS / IT and Data Science, Prof A K Bakhshi, Vice Chancellor, PDM University, said, "The CS /IT consists of different technical concepts such as programming languages, algorithm design, software engineering, computer-human interaction and the process of computation. Principal areas include database systems, networks, security, theory of informatics and bioinformatics. Data Analytics/Science is useful in studying internet users' behaviour and habits by gathering information from the users' internet traffic and search history."

Advantage of interdisciplinary programme

Going for an interdisciplinary programme has its own advantage, as the course offered by IIT Madras allows students opt for this program in their 3rd year which will earn them a bachelor's degree in their parent discipline (any of the engineering or biosciences degree IITM offers) and a Master's in data science. "The goal of this program is to train people with competency in a core discipline to become better data scientists," says Prof. Balaraman Ravindran, Head, Robert Bosch Centre for Data Science and Artificial Intelligence (RBC-DSAI) at IIT Madras.

There are several institutions offering B.Tech CSE with specialization in Data Science & Machine Learning to establish a strong foundation of Data Science & Machine Learning, based on agile methodology. The programmes' structure and curriculum design have been developed in association with the industry.

Demand for data analytics professionals

The role of Data Scientists has grown over 650% since 2012. As per an IBM report, the number of Data Science and Analytics job listings is projected to grow by nearly 364,000 listings to approximately 2,720,000 by 2020. Data Scientists are increasingly required to have skillsets beyond just statistical tools and techniques and achieving this requires a broader skill set across the data science and software engineering spectrum.

The McKinsey Global Institute a few years back stated that by 2018, the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know-how to use the analysis of Data Science to make effective decisions.

India is not far behind with 3344 Data Science & Machine Learning job openings on Indeed alone and over 4000 job postings on Naukri.com with salaries extending from Rs.3 lakhs p.a. to over 25 lakhs. "Data scientists are in high demand, with forecasts from IBM suggesting that the number of data scientists will reach 28 percent by 2020. In the US alone, the number of roles for all US data professionals will reach 2.7 million. Demand is clearly outstripping supply for data scientists," said Prof. Bakshi.

Job Roles

Data Scientist
Data Architect
Data Mining Engineer
Data Analyst
Data Engineer
Business Intelligence Analyst
Data Project Manager
Data Specialist
Data Developer

Career prospects

Jobs like 'Web Crawlers' and in-house 'Data Scientists' are nowadays needed in almost every industry, for purposes like collection of data from internal and publicly available data sources.

Adequate skills and experience are needed in industries for the tasks like the machine learning, computing and design of algorithms to create customized ETL tools to extract data from multiple sources/ IT platforms and to load them directly to data structures hosted in the cloud. Big Data technologies such as Hadoop', MongoDB, etc. are crucial for integrating both structured and unstructured data for clients to derive accurate insights from a given comprehensive set of data points.

Talking about the programme at IIT Hyderabad, Dr. Balasubramanian said that it will provide training on various aspects of AI, ranging from technical ones like machine learning algorithms to ethical issues as well. "It will make IIT Hyderabad students very sought after by the industry as their training will be comprehensive and well-rounded when working on smart AI-based technology solutions," he said.

Graduates interested in pursuing higher studies in the discipline may go for courses like M. Tech, MS, ME or PhD in India or abroad. Prof. Hanu Bhardwaj, HoD, CSE, Manav Rachna University, said, "Where do we see Data Science & Machine Learning in action in real-life? Pattern recognition in text, image, video or speech, making recommendations, making predictions in an online purchase, medical analysis from patient data etc."

Top recruiters

Successful graduates of the course are hired lucratively in areas such as Design Engineering, Computing/ IT Consulting and Solution Developing in capacities such as System/Network Administrators or IT Managers. This programme trains professionals to contribute to innovative hardware and software design of computer systems.

Some of the most popular firms that hire successful graduates of the specialization are Microsoft, Oracle, IBM, Infosys, TCS, HCL, Accenture etc., and in capacities such as Computer Engineer, Software Developer, Software Tester, Network Administrator, Database Administrator, Database Designer, Database Operator, Website Designer, Website Developer, Mobile Application Developer, Robotics Expert, Software Analyst etc.

Shutterstock

Industry-driven programmes

One of the advantages of these B.Tech courses is also that at most of the places they started either in collaboration with industry or industry has helped in designing of the curriculum, thus making it far more relevant. Several companies like Intel, IBM, Microsoft, Xebia have partnered with universities all over the country.

Manav Rachna University and Quantum University have started their B.Tech with Xebia where course and curriculum have been custom designed by data experts from Xebia who deal with such data engineering problems on a daily basis. "The students will have the opportunity of taking part in the annual hackathon, Internship support – Onsite/Remote, every student to be provided with 3 interview opportunities by Xebia, technical webinars for students and access to Xebia's Knowledge Sharing Sessions," said Prof Hanu Bhardwaj.

Mody University in Rajasthan, PDM University in Haryana, DIT University (Uttarakhand) have tied up with IBM where B.Tech in Computer Science & Engineering (Specialization in Big Data Analytics in association with IBM), B.Tech in Computer Science & Engineering (Specialization in Cloud Computing in association with IBM), is being offered. Great Lakes International University has started B.Tech in Big Data Analytics where students get a chance to study Data Science in analyzing data and subsequent presentation using data visualization techniques.

Prof. A K Bakshi,
Vice Chancellor, PDM University

Forecasts from IBM suggest that the number of data scientists will reach 28 percent by 2020. In the US alone, the number of roles for all US data professionals will reach 2.7 million. Demand is clearly outstripping supply for data scientists