

Data Science: The Science to understand Oil of 21st Century

Geetika Vashisht^{1*} and Vishesh Goyal²

¹Assistant Professor, Department of Computer Science, University of Delhi, India

²Student, Department of Computer Science, University of Delhi, India

***Corresponding Author:** Geetika Vashisht, Assistant Professor, Department of Computer Science, University of Delhi, India.

Received: October 18, 2021; **Published:** December 03, 2021

21st century is all about automation. Human race is obsessed with automation. Taking Amazon's shopping website as an instance, it uses customers' shopping data to figure out the shopping patterns and tries to predict your next order and to restock its warehouses nearest to you with what you are most likely to order so they can get products to you as soon as possible. This is called "prime now" under which they deliver 90 percent of things you need within 3 hours. Amazon even has patent called anticipatory shipping under which they will try to deliver things before you even order them based on your data. So, amazon is trying to become a little wishing machine you just open it and you get what you wish for. Also, apps like YouTube and instagram can keep you engaged for hours without even needing your input, they have this auto play feature under which you can passively enjoy for hours. So, what is driving all of these advancements? It's the data -the oil of 21st century and Data science is an ensemble of techniques to extract petrol from this oil which is in a sense useless in its raw form. Now, the question arises that where this data does comes from? It's from us, everything we do online all of our searches, likes, credit card swipes, Facebook posts everything is being collected into a trillion-dollar data industry. The reason why Facebook, Google, Amazon are the most powerful companies on planet is because last year data surpassed oil in value. These companies have mastered how to use the data they collect efficiently and reap huge number of profits from it. But let's take a step back and ask ourselves first!

What actually is Data Science: Now one of the most common definitions of data science is "Data science is a science which combines domain expertise, modeling tools from statistics and mathematics and computer algorithms to prepare model for a given problem using the data provided".

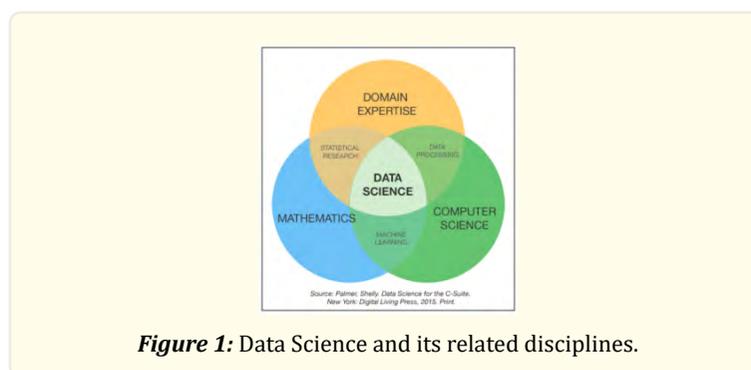


Figure 1 shows how data science is at the heart of so many other disciplines. It encompasses the use of analytics tools to notice patterns in data, connection between different parts of data. It uses statistical tools to make and test hypothesis. And then finally machine learning algorithms like as simple as linear regression or as advanced as deep learning techniques to make the most accurate model for a given problem. It tries to import a lot of things from already developed sciences. Now, because data science is such a new field, just

20-30 years old, there are no specific laws or theorems or principles like we have in other sciences. Many of them are either imported from other sciences or are very much problem specific such that it can't be generalized. So, you would ask if it imports so many things and in itself doesn't have many general laws and theorems then what is the use of Data Science!

Now, just because data science is new doesn't mean it's less impactful than other sciences like biology or physics. It is one of the most impactful sciences of 21st century. How? All of our devices from Alexa to Google assistant to auto-captioning in YouTube, all are the products of data science. Data science is used to devise advertisement techniques for companies so as to gain maximum profit, it's also used in recommendation systems of Amazon I bet all of you have seen that tab by these things together or similar products due to which you shop for double the amount you originally intended. Data science is used to make offers personalized just for you by keeping your shopping habits in mind, for example you might be fan of soft-drinks so here buy more so as to get discount custom made just for you. Data science is today revolutionizing every field weather its advertising, economy, defence sector of a country and even helping traditional sciences like biology, to move forward. So much so the job title of a Data scientist is most in-demand right now.

Is data-science hyped too much: Now, one of the most common doubt is, isn't this science hyped too much, is it really going to deliver the results it promises. Well, there is no clear-cut straight forward answer. It is new and in its early stages only. It could be the next dot com bubble burst which can cause markets to crash or could create next Facebook and google. But no one really knows the answer. What everyone knows is that companies and countries are investing a lot of money in it. Every country, company wants a good share of this new industry and that's why this field is attracting investments worth trillions of dollars. Today we are the forefront of Fourth industrial revolution which is going to be the biggest and fastest revolution humanity has ever seen and it is not possible without data-science. Like mechanical physics was basis for first industrial revolution in which industry became the backbone of society, then second industrial revolution was driven by thermal physics with the advent of internal combustion engine and new sources of energy like electricity and light were discovered similarly data science will be the basis for fourth industrial revolution in which new source of as I will quote "energy" will be "Data".

We are today in what is called the information age. This age is about how wisely and accurately any company or country uses data if it wants to reap benefits from this age, and none of this is possible without data science. Data in itself is like an ore which without proper extraction is just an alien language which is of no use.

Volume 1 Issue 3 December 2021

© All rights are reserved by Geetika Vashisht., et al.