



## ETA NEWS BULLETIN

Electronics and Telecommunication Engineers Association  
An Electronics and Telecommunication Engineering Department Publication

VOLUME 13:ISSUE 1

SEPTEMBER 2017

### WHAT'S INSIDE

#### PAGE 2

Automation and benefits of Data Science

#### PAGE 3

Into the world of Data Science

#### PAGE 4

Amusement with facts, crossword and quiz

## DATA SCIENCE

*Have you ever looked for shoes online and found related advertisements on facebook and other website continuously for a week ? Or say, you are chatting with your best friend and the keyboard suggests you the exact words that you want to use in your sentence? How does YouTube show all your favorite videos on your homepage? Well, these are answered by Data Science.*

Data science, also called as data-driven science is an inter-disciplinary field about scientific methods, processes, and systems to extract knowledge or insights from data in various forms either structured or unstructured, similar to data mining data science is a "concept to unify statistics, data analysis and their related methods" in order to "understand and analyze actual phenomenon" with data. It employs techniques and theories drawn from many fields within the broad areas of mathematics, statistics, information science and computer science, in particular from the sub-domains of machine learning, classification, cluster analysis, data mining, databases and visualization. Data science is ultimately about using the data in creative ways to generate business value. The goal of the Internet of Things (IoT) is to acquire data from various embedded systems and impart analytical processes on that data to improve performance, efficiency, and business outcomes. The ability to create analytics that process massive amounts of business and engineering data is enabling designers in many industries to develop intelligent products and services.

Following levels are invoked:



**A) DATA SCIENCE : DISCOVERY OF DATA INSIGHT** It is all about uncovering findings from the data .It's about surfacing hidden insight that can help enable companies to make smarter business decisions. For example, NETFLIX data mine movie viewing patterns to understand what drives user interest and uses that to make decisions on which NETFLIX original series to produce .TARGET identifies what are major customer segments within its base and the unique shopping behaviors within those segments, which helps to guide message to different market audiences. PROCTOR and GAMBLE utilizes time series models to more clearly understand future demand, which help plan for production levels more optimally.

**B) DATA SCIENCE: DEVELOPMENT OF DATA PRODUCT** - A "data product" is a technical asset that: (1) Utilizes data as an input, and (2) Processes that data to return algorithmically-generated results. The classic example is the RECOMMENDATION ENGINE, which ingests user data and makes personalized recommendations based on the data. Here are some examples: AMAZON'S recommendation engines suggest items for you to buy, determined by their algorithms.NETFLIX recommends movies to you.SPOTIFY recommends music to you. GMAIL spam filter is a data product - an algorithm behind the scenes processes incoming mail and recommends if a message is junk or not. COMPUTER VISION- used for self-driving cars is also a data product: the machine learning algorithms are able to recognize traffic lights, other cars on the road, pedestrians.

This is different from the "DATA INSIGHTS" section, where the outcome to is to provide advice to an executive to make a smarter business

decision. In contrast, a **DATA PRODUCT** is technical functionality that encapsulates an algorithm and is designed to integrate directly into core applications.

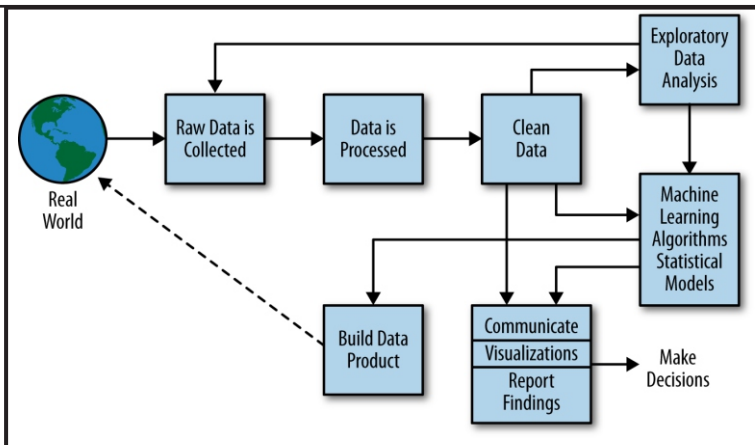
#### AUTOMATION OF DATA SCIENCE:

The biggest knock against the data science is that it's extremely labor-intensive.

Market watcher Gartner says that's about to change, predicting that by 2020, fully 40% of all data science tasks will be automated. According to Gartner, software vendors in the data management and analytics spaces are building automation features into their platforms. They are placing particular emphasis on automation processes for data integration and data preparation.

#### BENEFITS

In some implementations, analytics are performed in the cloud with the intent to improve the performance of existing embedded



systems. In other cases, analytics are better run directly in an embedded system. For instance, a design team at Scandia, the Swedish truck manufacturer, embeds analytics into their emergency braking systems to provide real-time crash avoidance to reduce accidents and meet stringent EU safety regulations.

Engineering data from cameras and radar are processed in real time for object and road marking detection, which is subsequently fused to signal collision warning alerts and automatic brake requests. The accelerating Internet of Things trend towards smarter and more connected sensor networks is only adding to that pressure. This has the benefit of shrinking the amount of data that is transferred over the network, which reduces the cost of transmission and can lower the power consumption of wireless devices.

Ref: <https://datajobs.com/what-is-data-science>

## ETA TEAM



STAFF INCHARGE	Prof. Ashwini Katkar
SECRETARY	Uzmanaz Shaikh
JOINT SECRETARY	Deepak Gavandi
TREASURER	Prathamesh Kadam
EDITING TEAM	Sakshi Naik, Rashmi Patil T.E. : Shrishti Shetty, Kalaksha Gunera Sanchit Akre, Jahnvi Mehta S.E. :Deeksha Kamath, Umang Waghela
ETA WALL TEAM	Prathamesh Gupta T.E. :Amit Kanswal, Shubham Tiwari S.E. :Pratik Waghela, Saili Sakpal
DESIGNING TEAM	Dheeraj Singh, Aditi Dave T.E. :Rajas Patil, Pankaj Gidwani S.E. :Omkar Samant, Krutik Sankhe
SPONSORSHIP TEAM	Alkesh Khedekar T.E. :Sagar Agrawal S.E.:Vibha Parab, Viraj Rasal
PUBLICITY TEAM	Pratik Pandit T.E. :Pratik Jadhav S.E. :Sakshi Dhanmeher, Ashwini Rathod

## ACADEMIC BRILLIANCE

ETA congratulates the following students for achieving great heights in their classes in the university exam held in May 2017.

#### **BE (2016-17)**

TOPPERS NAME	CGPI	RANK
Nimaje Someshwar Ankush	9.05	1st
Surve Bhakti Prakash	8.86	2nd
Yadav Neelam Rajendraprasad	8.86	3rd

#### **TE (2016-17)**

TOPPERS NAME	V SEM SGPI	VI SEM SGPI	AVG	RANK
Naik Bhavesh	9.11	9.5	9.30	1st
Baurai Anurag	9.56	8.86	9.21	2nd
Shetty Nishita	9.34	8.86	9.1	3rd

#### **SE (2016-17)**

TOPPERS NAME	V SEM SGPI	VI SEM SGPI	AVG	RANK
Negi Pratik	9.38	10	9.69	1st
Patelia Vrutika	9.23	9.14	9.18	2nd
Tiwari Shubham	8.62	9.46	9.04	3rd

**CONGRATULATIONS !!!!**

# INTO THE WORLD OF DATA SCIENCE

Data science as a blend of three major areas-Mathematics expertise,Technology/Hacking skills and business,strategy accumen proves to be at the heart of data mining insights and building out product which will add up to the market value. Apart from empowering the management,it identifies the opportunities to make the best use of data available.

Witnessing the increasing scope and need of data science, we have our VCET 2010 ALUMNUS **Ankur Tiwari**, who have been gaining sound knowledge about the various aspects of data science.Further illuminating his career,he completed his MBA(finance) in the year 2014.He is a journalist with interest in the fields of technology, design thinking, new product development and branding.

Having previously worked at **Embecon systems** (<http://embecon.com/>) as an embedded system engineer and **L&T Finance** (<https://www.ltf.com/>) as Project manager in financial projects, he is currently involved in the following setups -

1.Thoughtlytics Internet Pvt. Ltd. (Cofounder) An internet startup developing proprietary algorithms in the field of data science and machine learning with an aim of providing 'Thoughtful analytics'.

2.Intuitive Embedded Tech (Cofounder) A tech education setup providing interactive education in the fields of embedded systems, robotics, IOT, Python and C programming, and Data science.

Visualising his precious experiences, we would like to request him to share some of his knowledge with us.Here,we have some important and interesting questions being answered by him as interviewed by Rajas Patil and Kalaksha Gunera.

## Q.What can be a brief view about Data Science according to you?

Experience they say, is the only source of knowledge, which allows a person to take better decisions. An inquisitive mind must ask, "What is that most crucial attribute that experience provide, which makes one more knowledgeable than not so experienced?" Applied researchers and scientist have pointed towards the vast amount of data stored in the neurons of the mind of an experienced person towards their superior decision making capabilities. Data about people, places, situations, behaviours and much more, stored in our brain, unconsciously help us to make better decisions.

Data science is the field of enabling machines to make smart decisions based on defined datasets and dedicated algorithms. It covers the wide range of topics and involves a dynamic cross functional team to create successful products. Key steps involved in Data science are

- 1.Data collection
- 2.Data integration and cleaning
- 3.Machine learning

## Q.What would you like to say about 'Winning over the Jargons: Data vs Big Data'?

Consider the case of a simple website which asks its users to register using email id, full name, gender, age and password. If the website has 10,000 registered users, then it will have 10000 data sets with 5 data points in each. Not much can be inferred about a user from these data points except the known details of age, gender, name and email.

Now consider the case of Google, which has these primary details of all its users (via Gmail), and has built up a huge database of all the searches you have made in last 10 years or so. It has in its record among many other things the terms you had searched for, the places you did visit and the food you had ordered. Now add the data points of all the videos you had watched on YouTube in this data (data points in a video can be 'category', 'sub-category', 'tags', 'length', 'fps', 'channel region' etc) and Google know more about you than anyone else in this world. With over a billion people using its services, Google has billion data sets with thousands of data points in each. This kind of high volume and wide variety of data is called Big Data.

## Q.What are the main reasons for Big Data becoming such a buzzword?

There are two main reasons for Big Data becoming such a buzz word:-

1.Size of the data is itself a problem.We are talking about gigabytes to petabytes of the data. Traditional systems of data management and techniques of working with data do not work here anymore. The computational challenges involved in working with the data of this size have given birth to altogether new set of systems.

2.The information stored in this size and variety of data is beyond anything we had before. Patterns and relations that can be found within this data have huge commercial value. Machines for the first time can realistically mimic human intelligence and if experts to be believed, can surpass it by 2045.A

## Q.What is takes to be a good data scientist?

A data scientist needs to be a good statistician, programmer and product person. He/she should be able to analyse the data, test various algorithms on that data by writing programs to find meaningful insights and should have product knowledge to be able to decide which insights are actionable and which are not.

For a beginner, expertise in Microsoft excel, non-relational databases and Python programming are must.

## Q.What are the applications wherein Data Science is being used?

Almost all the internet products that we use today, use data science in one way or other. Giants like Google, YouTube, Facebook, Amazon LinkedIn and Netflix (among many others) have created excellent data driven products that are being used by billions of people every day. A mention of few known data driven products will help in understanding the potential of data science.

**a.Speech recognition:** Google has made significant progress by collecting voice data and processing those using proprietary algorithms. They have been able to integrate voice search with their search engine.

**b.Product recommendations:** Amazon uses purchase and browsing history of its users to successfully recommend products to them every time they visit the site.

**c.Facebook and LinkedIn** use data science to suggest people you may know and would like to add as friend.

**d.Netflix and YouTube** recommends videos using their data science capabilities.

## Q.Would you like to say something more about Data Scientists?

For a data scientist, the jeans you buy, friend's profile you visit, terms you search and the video you watch on different internet platforms are the 'data points' which enable him to create a virtual personality of yours and provide a customised service with great accuracy.

## Q.What would you like to comment on 'Data science is sensor driven world'?

It's a matter of time before a network of all kind of sensors will occupy our public places generating tremendous amount of data. With all this data being stored on clouds, next gold rush is going to be making sense of this data. Data scientists who can make useful and commercially viable products using this data, will be much sought after.

-As interviewed by Rajas Patil,Shrishti Shetty.

(Ankur Tiwari can be reached [www.ankurt.04@gmail.com](mailto:www.ankurt.04@gmail.com))

