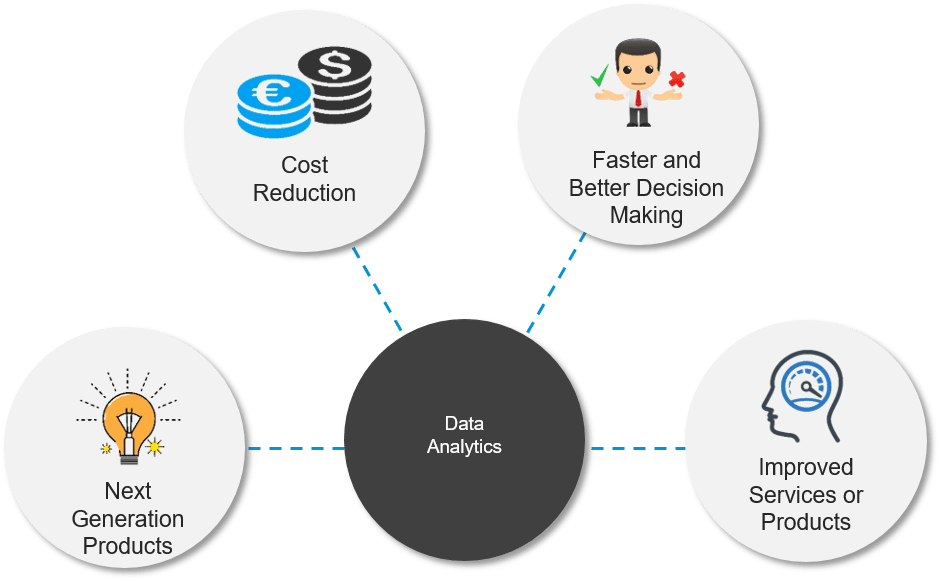
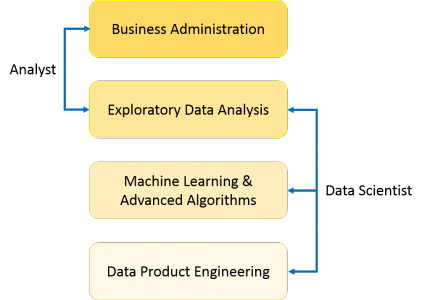
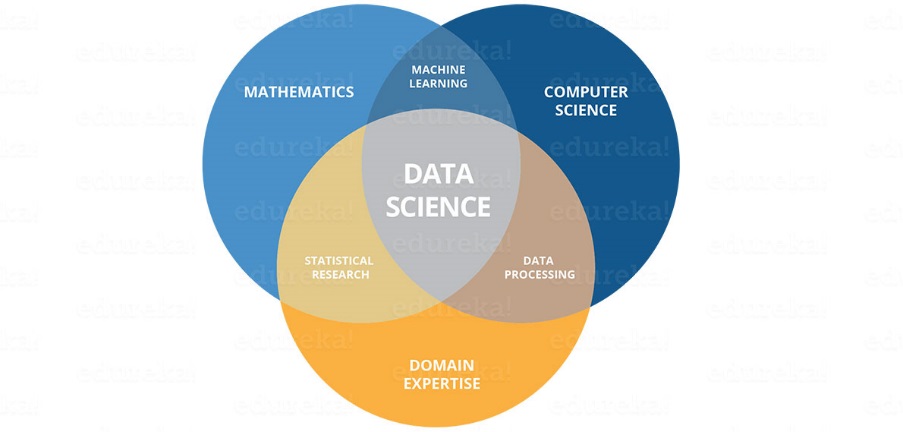
**Why Data Science?**



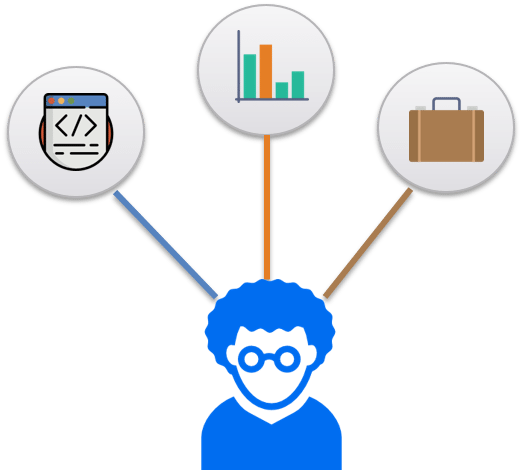
**What is Data Science (Data Driven Science)?**

* Data science makes use of scientific methods, processes and systems to extract knowledge or insights from data in various forms, i.e either structured or unstructured.





**Who is data scientist?**

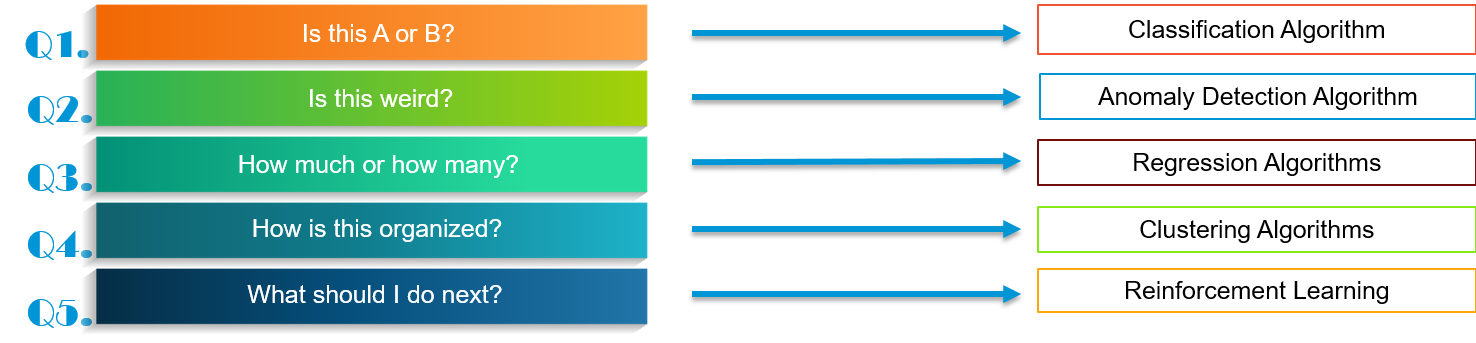


**Skills required:**

1. Be very innovative and distinctive in his approach in applying various techniques intelligently to extract data and get useful insights in solving business problems and challenges.
2. Have the ability to locate and construe rich data sources.
3. Have a hands-on experience in Data mining techniques such as graph analysis, pattern detection, decision trees, clustering or statistical analysis.
4. Develop operational models, systems and tools by applying experimental and iterative methods and techniques.
5. Analyze data from a variety of sources and perspectives and find out hidden insights.
6. Perform Data Conditioning – that is, converting data into a useful form by applying statistical, mathematical tools and predictive analysis.
7. Research, analyze, execute, and present statistical methods to gain practical insights.
8. Manage large amounts of data even during hardware, software and bandwidth limitations.
9. Create visualizations that will help anyone understand the trends in data analysis with ease.
10. Be a team leader and communicate effectively with other business analysts, product Managers and Engineers.

**How to solve a problem in Data sceince?**

* Problems in Data Science are solved using Algorithms. But, the biggest thing to judge is which algorithm to use and when to use it?
* Basically there are 5 kinds of problems which you can face in data science.



**Is this A or B?**

* It is an categorical problem, the answer could be either 0 or 1 , Yes or No and True or False.
* Example:

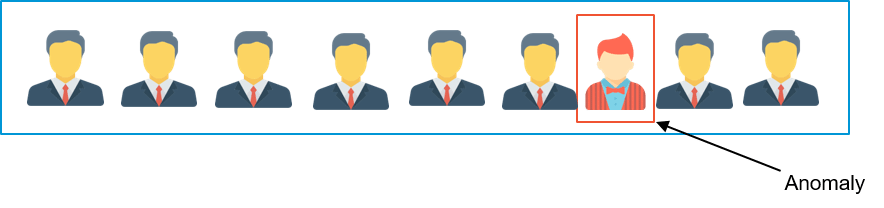
1. Do we have session today?

Yes or No

1. Tickets available or not?

Yes or No

**Is this weired?**

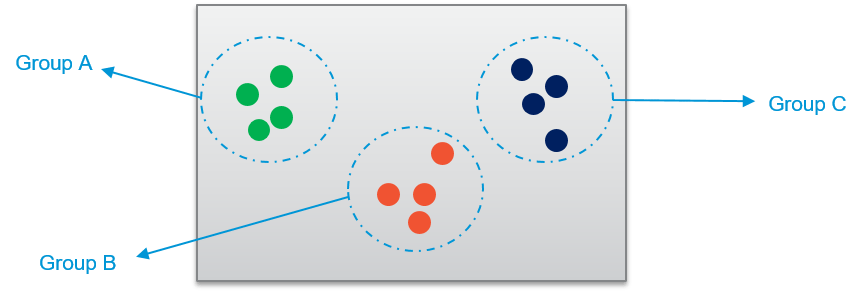
* These kind of problems deals with the patterns.
* As you can see in the below image , you can say the red guy is looking different from others. Whenever there is a break in pattern or usual behaviours it rasie the flag that particular event and make it review for us. So that we can easily figure out the odd one out.
* Example:
  + Credit card fraud detection.
  + Implementing security and reducing human efforts on surveilance.

**How much or How many?**

* Whenever we deal with numbers we solve it using regression algorithms.
* Example:
  + What will be the temparature tommorow? 

**How is this Organized?**

* Say you have some data, you don’t have any idea about the data,how you make sense out of this data?
  + You can solve it using **clustering** **algorithms**. Clustering algorithms group the data based on the common characteristics.
  + As you can see in the below image the data are seperated based on the some attributes or behaviours.
  + For example, seperating the customers into Platinum, gold and silver customer based on the customer purchase.

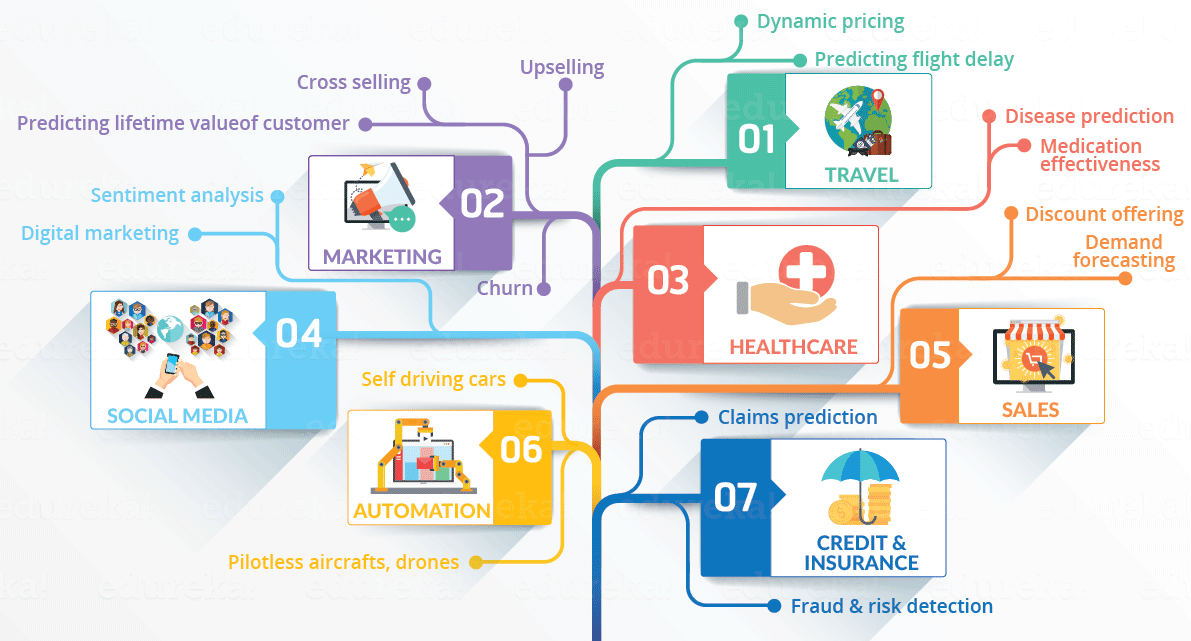


**What should I do Next?**

Whenever you want to take any decision based on the past data or training that we have given, it involves the Reinforcements Algorithms.

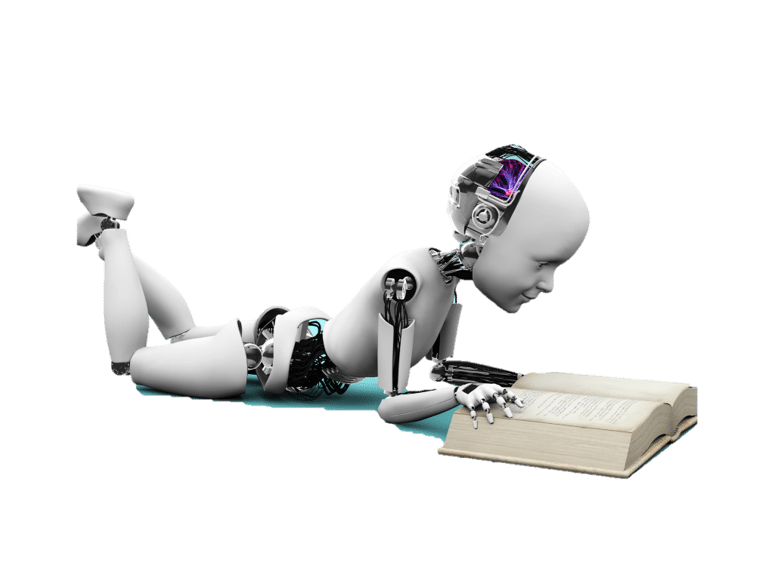
**How do these Algorithms work?**

* + - Based in human sychology.
    - Rather than teaching computers, you let it decide what to do and end of the action you give either positive or negative feedback.
    - It is just like training your dog,You cannot control what your dog does,right?, But you can scold him when he commit mistakes. Similarly you can patting him on the back when he does what I expected.
    - For example, lets imagine you are training temparature control system so whenever no of people increases in the room either the temparature need to be increased or decreased. Since our system doesn’t undestand anything, it takes random decision. Let’s suppose it increases the temparature , then you give the negative feedback. With this the system will understand whenever the number of people increases it has to decrease the temparature, never increase the temparature.



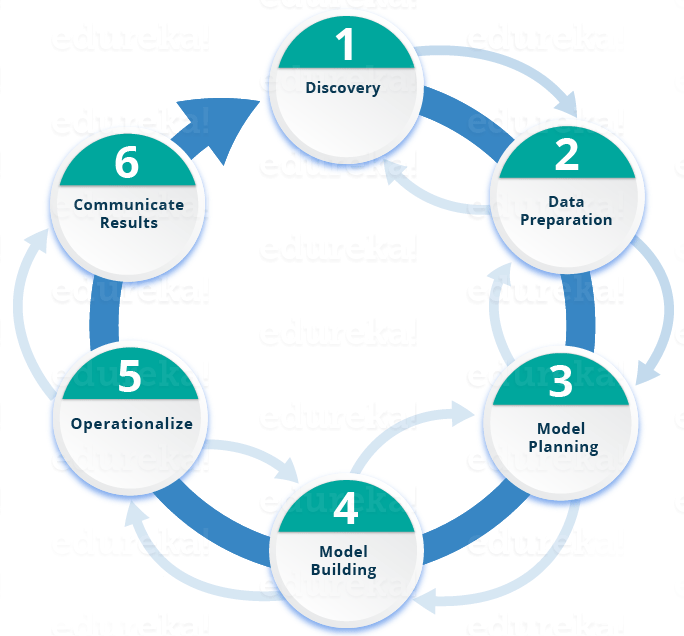
**Machine Learning :**

For all above actions need to be completed we always saying Algorithms right? So what is machine learning algorithms?



* Its is a type of artificial intelligence that makes the computers capable of learning on their own.(ie) without explicitly being programmed.

**Life cycle of Data Science:**



**Business Intelligence (BI) vs. Data Science**

* BI basically analyzes the previous data to find hindsight and insight to describe the business trends. BI enables you to take data from external and internal sources, prepare it, run queries on it and create dashboards to answer the questions like quarterly revenue analysis or business problems. BI can evaluate the impact of certain events in the near future.
* Data Science is a more forward-looking approach, an exploratory way with the focus on analyzing the past or current data and predicting the future outcomes with the aim of making informed decisions. It answers the open-ended questions as to “what” and “how” events occur.

|  |  |  |
| --- | --- | --- |
| **Features** | **Business Intelligence (BI)** | **Data Science** |
| Data Sources | Structured (Usually SQL, often Data Warehouse) | Both Structured and Unstructured  ( logs, cloud data, SQL, NoSQL, text) |
| Approach | Statistics and Visualization | Statistics, Machine Learning, Graph Analysis, Neuro- linguistic Programming (NLP) |
| Focus | Past and Present | Present and Future |
| Tools | Pentaho, Microsoft BI, QlikView, R | RapidMiner, BigML, Weka, R |

**Data Science Application:**

* [Fraud and Risk Detection](https://www.edureka.co/blog/data-science-applications/#fraudandriskdetection)
* [Healthcare](https://www.edureka.co/blog/data-science-applications/#healthcareindatascience)
* [Internet Search](https://www.edureka.co/blog/data-science-applications/#internetsearch)
* [Targeted Advertising](https://www.edureka.co/blog/data-science-applications/#targetedadvertising)
* [Website Recommendations](https://www.edureka.co/blog/data-science-applications/#websiterecommendations)
* [Advanced Image Recognition](https://www.edureka.co/blog/data-science-applications/#imagerecognition)
* [Speech Recognition](https://www.edureka.co/blog/data-science-applications/#speechrecognition)
* [Airline Route Planning](https://www.edureka.co/blog/data-science-applications/#airline)
* [Gaming](https://www.edureka.co/blog/data-science-applications/#gaming)
* [Augmented Reality](https://www.edureka.co/blog/data-science-applications/#augreality)