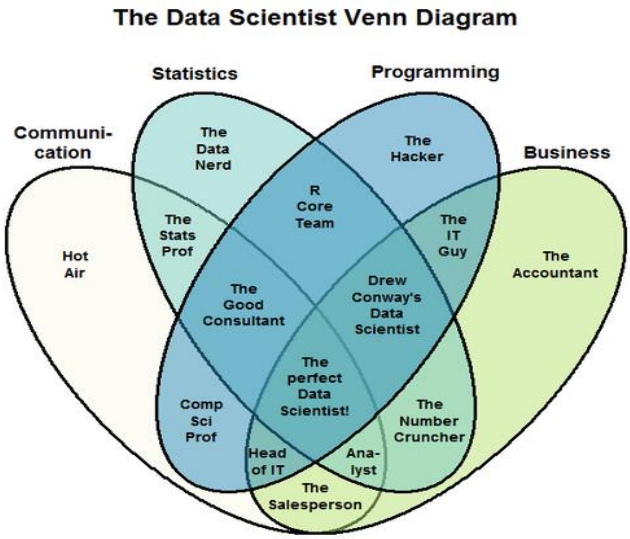
BITS AND BYTES
bits.and.bytes.8b1b@gmail.com

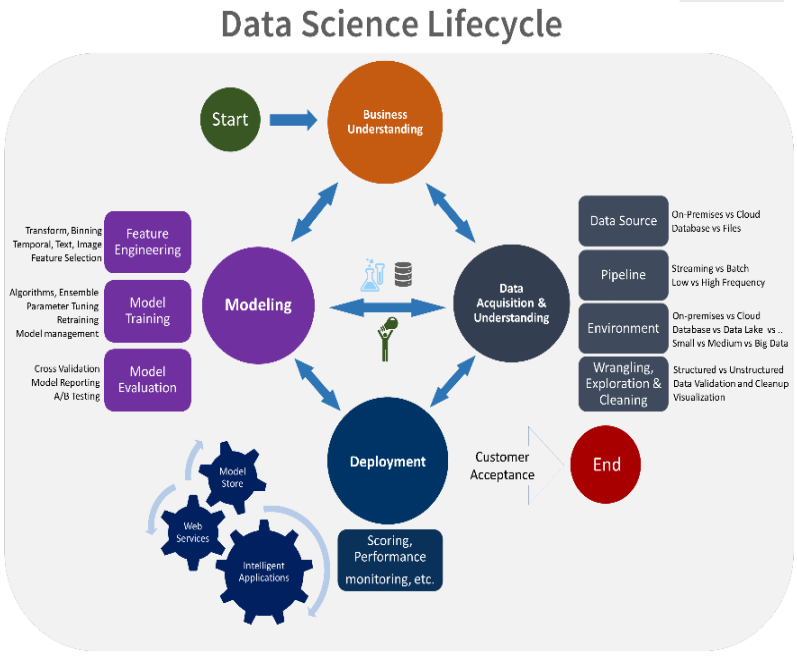
INTRODUCTION TO DATA SCIENCE

Data science is an inter-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structured and unstructured data. Data science is related to data mining, machine learning and big data.

Data science enables us to explore vast amount of data and gain insights from it. With proper filtering and preprocessing, data becomes suitable for extracting information.

Data science is a "concept to unify statistics and data analysis and their related methods" in order to "understand and analyze actual phenomena"





with data. It uses techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, domain knowledge and information science.

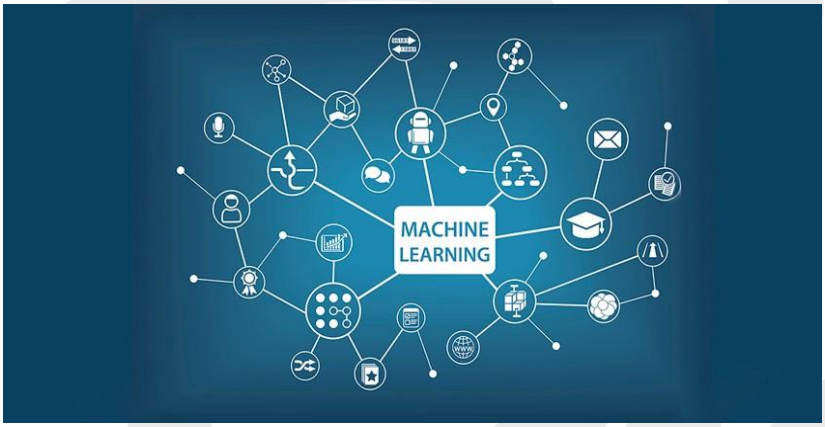
Data science is used almost everywhere. The job of a data scientist will require sophisticated skills to extract information from the data.

Now, let's discuss about data mining, machine learning and big data.

Data Mining

Data mining is a very trending topic in today's world. Companies turn raw data into useful information using data mining. Business strategies can be made based on the customer information analysis. Production can be made cost efficient using data mining.

Also companies use it to predict future cost, credit card fraud detection, spam email filtering and many more.



Machine Learning

Machine Learning is very widely used in the industries. It is very popular subject and widely studied. Various machine learning models are out there that can perform many tasks with very high precision. We will discuss about this in the next module.

Big Data

As its name suggests, big data implies large amount of data gathered in one place. This may be both structured and unstructured data. Some example of big data are, social media database, stock exchange information etc. Using algorithms, data scientists try to gain insight from this humongous data.

This concludes our today's session. Feel free to ask any query in the comment section or in the mailing address bits.and.bytes.8b1b@gmail.com. For any other issues, you can also use this same email address.

References:

1. https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.datanami.com%2F2018%2F09%2F17%2Fimproving-your-odds-with-data-science-hiring%2F&psig=AOvVaw3l1TCoZFcM3PUrNiBZt74h&ust=1604924163589000&source=images&cd=vfe&ved=0CAMQjB1qFwoTCMDd48v28uwCFQAAAAAdAAAAABAJ
2. https://images.squarespace-cdn.com/content/v1/5ccb715016b640627a1c2782/1586907293788-T2DQBYUM8QN1HFIGIR5G/ke17ZwdGBToddI8pDm48kOyctPanBqSdf7WQMpY1FsRZw-zPPgdn4jUwVcJE1ZvWQUxwkmyExglNqGp0IvTJZUJFbgE-7XRK3dMEBRBhUpyD4IQ_uEhoqbBUjTJFcqKvko9JlUzuVmtjr1UPhOA5qkTLSJODyitRxw8OQt1oetw/data-science-venn-diagram-stephan-kolassa.png?format=500w
3. https://www.google.com/url?sa=i&url=https%3A%2F%2Fdocs.microsoft.com%2Fen-us%2Fazure%2Fmachine-learning%2Fteam-data-science-process%2Flifecycle-business-understanding&psig=AOvVaw3l1TCoZFcM3PUrNiBZt74h&ust=1604924163589000&source=images&cd=vfe&ved=0CAMQjB1qFwoTCMC70OD38uwCFQAAAAAdAAAAABAd
4. https://www.google.com/url?sa=i&url=https%3A%2F%2Fdata-flair.training%2Fblogs%2Fdata-mining-tutorial%2F&psig=AOvVaw1sjdhO98QSM15nWyI7vWPr&ust=1604924854361000&source=images&cd=vfe&ved=0CAMQjB1qFwoTCMiMoZL58uwCFQAAAAAdAAAAABAJ
5. https://www.google.com/url?sa=i&url=https%3A%2F%2Ftowardsdatascience.com%2Fmachine-learning-for-beginners-d247a9420dab&psig=AOvVaw1Vhg0LvLJrK8bt3N7UpnRv&ust=1604925673738000&source=images&cd=vfe&ved=0CAMQjB1qFwoTCNCt85j88uwCFQAAAAAdAAAAABAJ
6. https://www.google.com/url?sa=i&url=https%3A%2F%2Ftowardsdatascience.com%2Fwhat-is-big-data-lets-answer-this-question-933b94709caf&psig=AOvVaw21dX3W1GhwUETTNM7c62UD&ust=1604926066150000&source=images&cd=vfe&ved=0CAMQjB1qFwoTCIjn5tv98uwCFQAAAAAdAAAAABAD