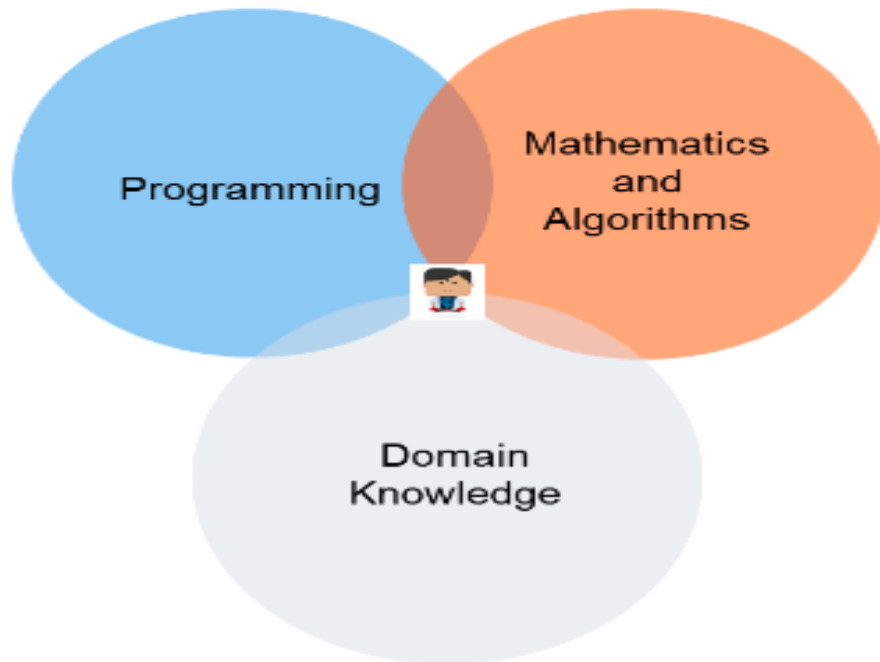


1. What is Data Science?

Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract or extrapolate knowledge and insights from noisy, structured and unstructured data, and apply knowledge from data across a broad range of application domains

Data science is the field of study that combines domain expertise, programming skills, and knowledge of mathematics and statistics to extract meaningful insights from data.



2. What is Artificial Intelligence?

Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems. Specific applications of AI include expert systems, natural language processing, speech recognition and machine vision.

3 Types of Artificial Intelligence

Artificial Narrow Intelligence (ANI)



Stage-1

Machine Learning

- Specialises in one area and solves one problem



Siri



Alexa



Cortana

Artificial General Intelligence (AGI)



Stage-2

Machine Intelligence

- Refers to a computer that is as smart as a human across the board

Artificial Super Intelligence (ASI)



Stage-3

Machine Consciousness

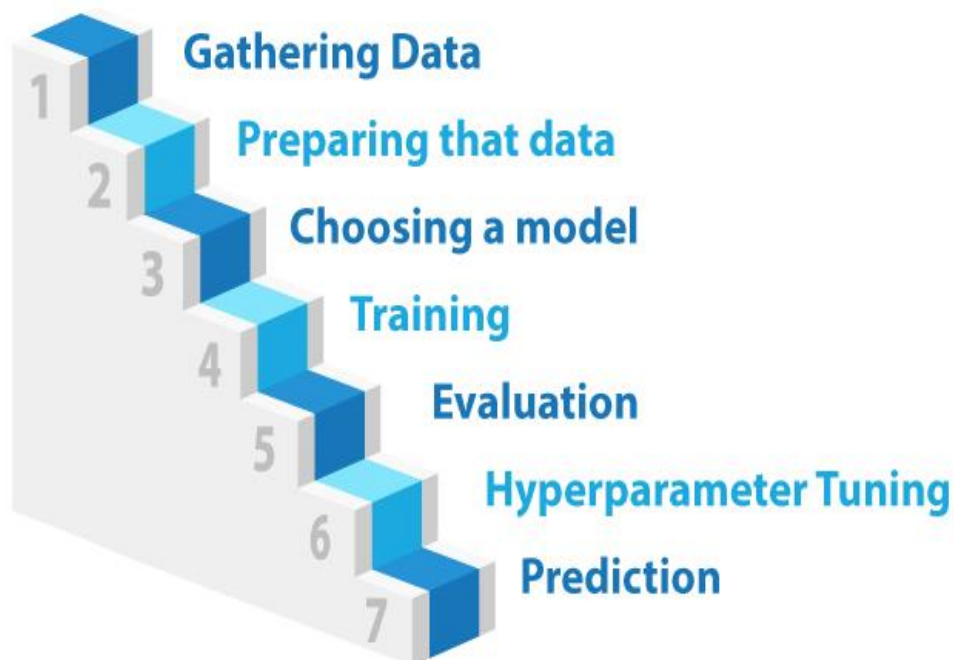
- An intellect that is much smarter than the best human brains in practically every field

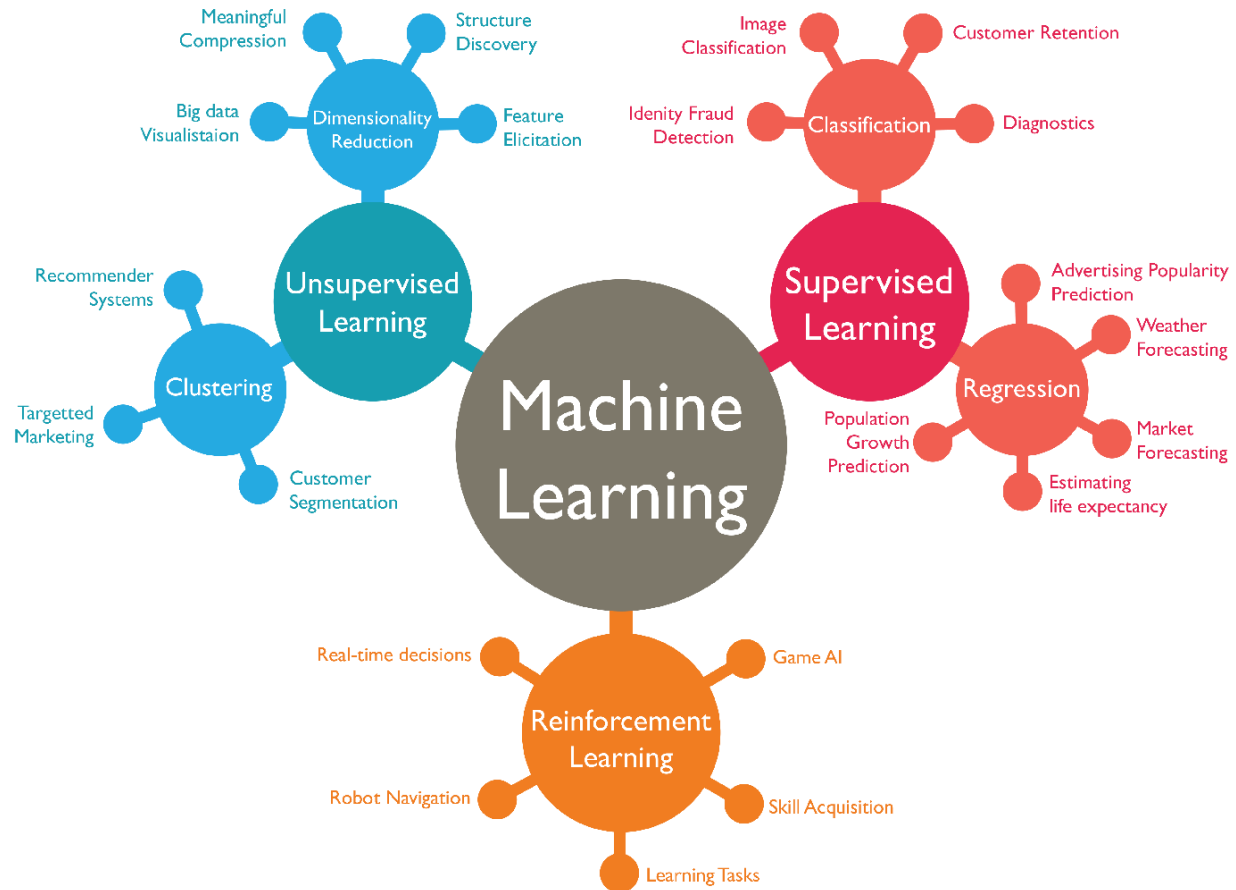
3. What is Machine Learning?

Machine learning is a subfield of artificial intelligence, which is broadly defined as the capability of a machine to imitate intelligent human behavior.

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

7 steps of Machine Learning





4. Future of Data Science and Real-life applications.

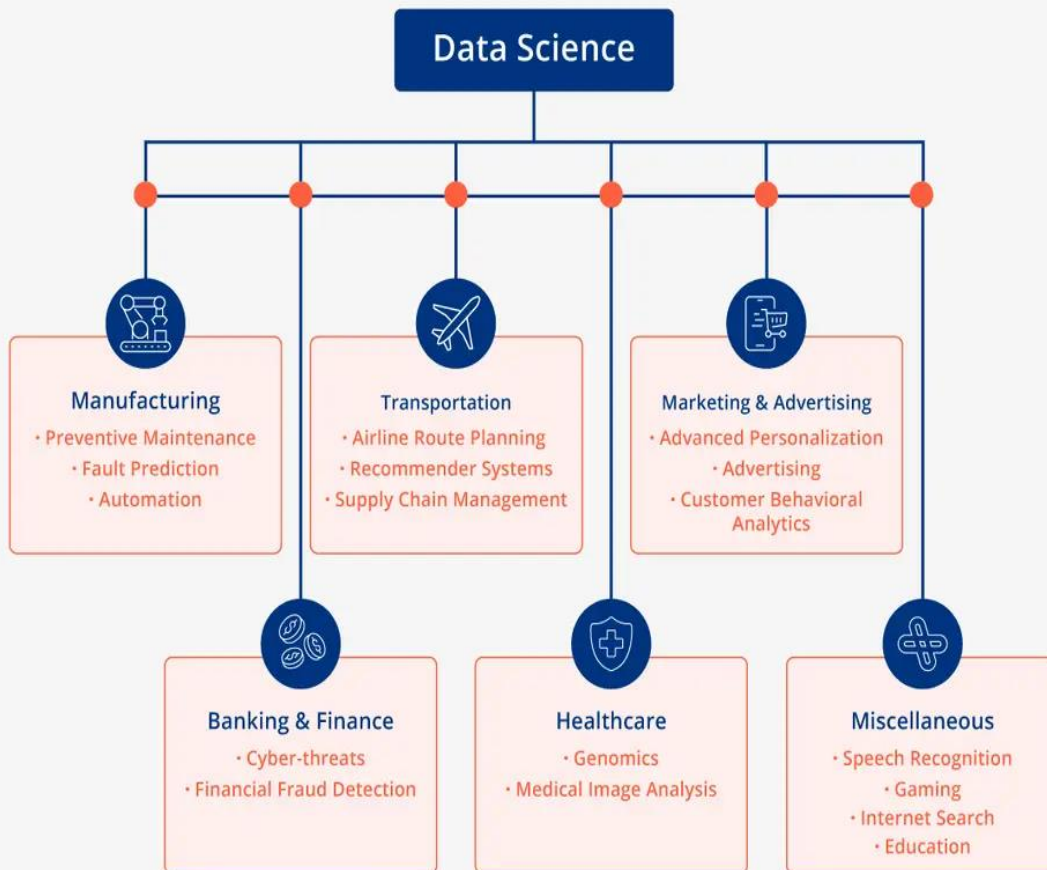
You can think about the data increase from IoT or from social data at the edge. If we look a little bit more ahead, the US Bureau of Labor Statistics predicts that by 2026—so around six years from now—there will be 11.5 million jobs in data science and analytics.

How data science used in real life?

Healthcare: Data science can identify and predict disease and personalize healthcare recommendations.

Applications of Data Science

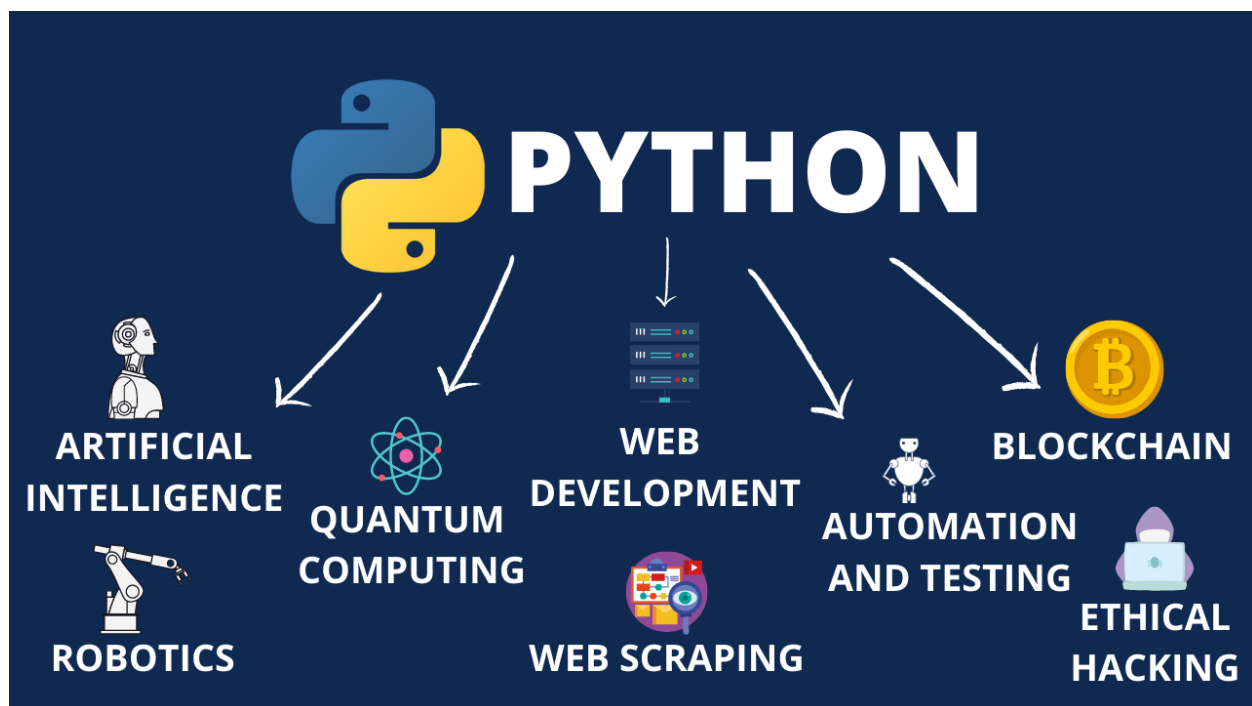
Data Science has disrupted every industry. Faster Computing, Cheaper Storage and Powerful Data Science Tools reduce the time to market for consumer products and better our Quality of Life.



5. Why Python so popular? Future of Python.

The python language is one of the most accessible programming languages available because it has simplified syntax and not complicated, which gives more emphasis on natural language. Due to its ease of learning and usage, python codes can be easily written and executed much faster than other programming languages.

Python is often employed to support machine learning.



What is the future of Python?

The future of Python is thus set in stone as the go-to language for developers engaging in deep machine learning projects, companies looking to scale up by using big data analytics or efforts aimed at achieving social applications of automation!

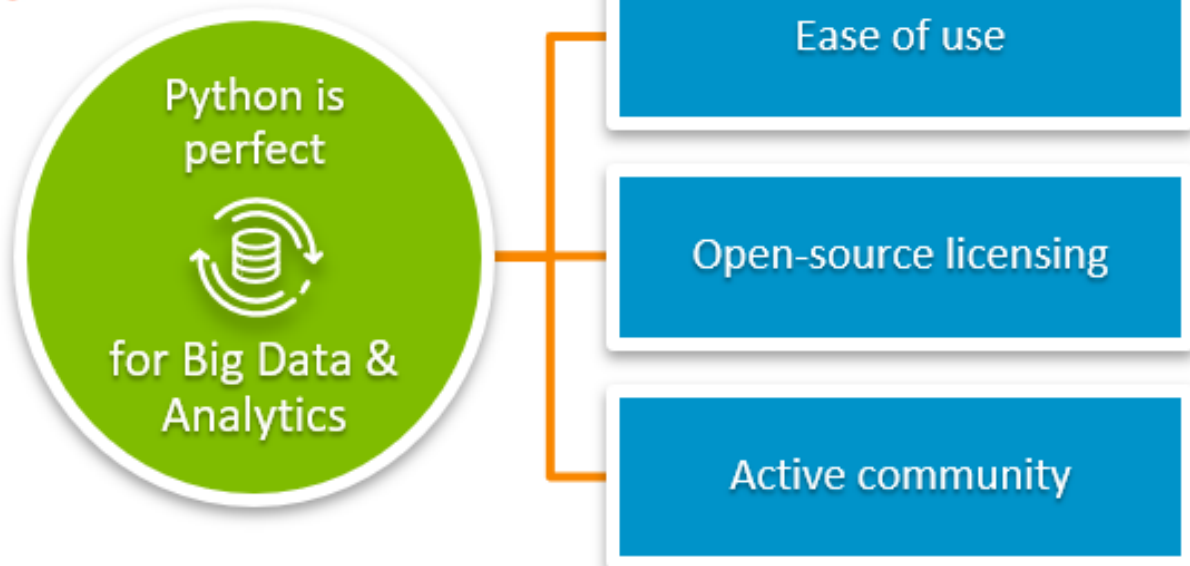
Signs Why Python is the Programming Language of the Future



6. Why Python for data analytics?

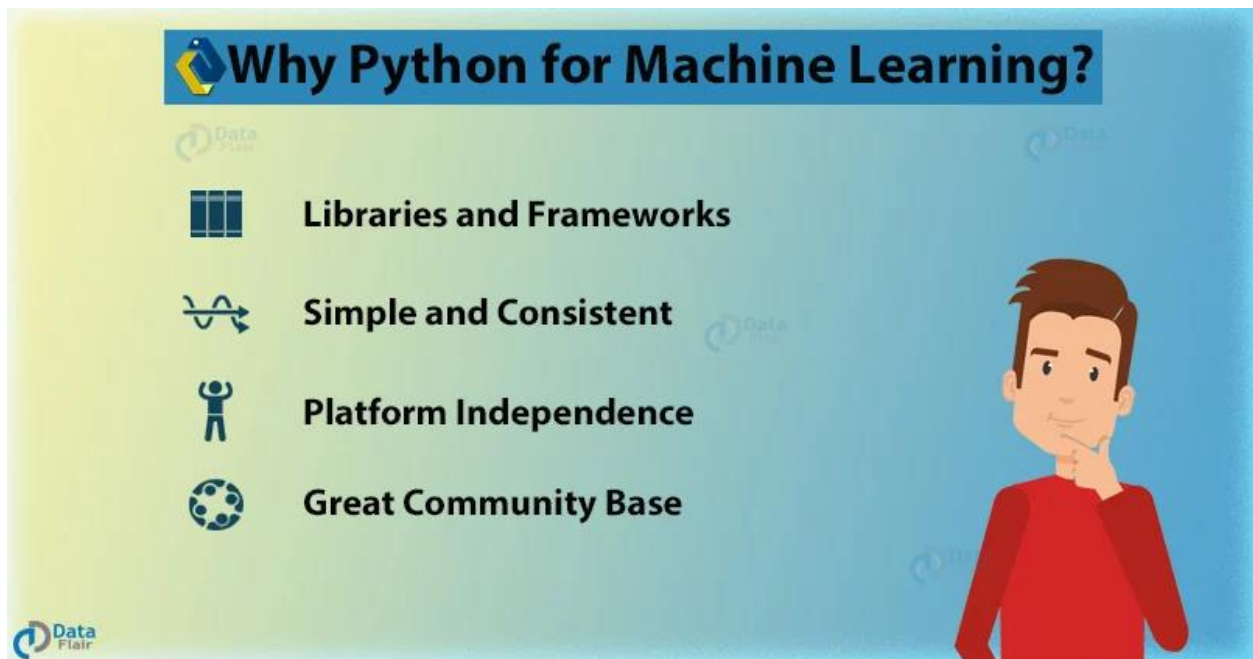
Thanks to Python's focus on simplicity and readability, it boasts a gradual and relatively low learning curve. This ease of learning makes Python an ideal tool for beginning programmers. Python offers programmers the advantage of using fewer lines of code to accomplish tasks than one needs when using older languages.

- It's Flexible
- It's Easy to Learn
- It's Easy to Learn
- It's Well-Supported



7. Importance of Python in machine learning.

Python offers all the skillsets that are required for a machine learning or AI project – stability, flexibility and a large number of tools. Python helps developers to be productive and confident about the product that they are manufacturing, from the stages of development to deployment and till the maintenance stage.



8. Importance of Python in deep learning & NLP.

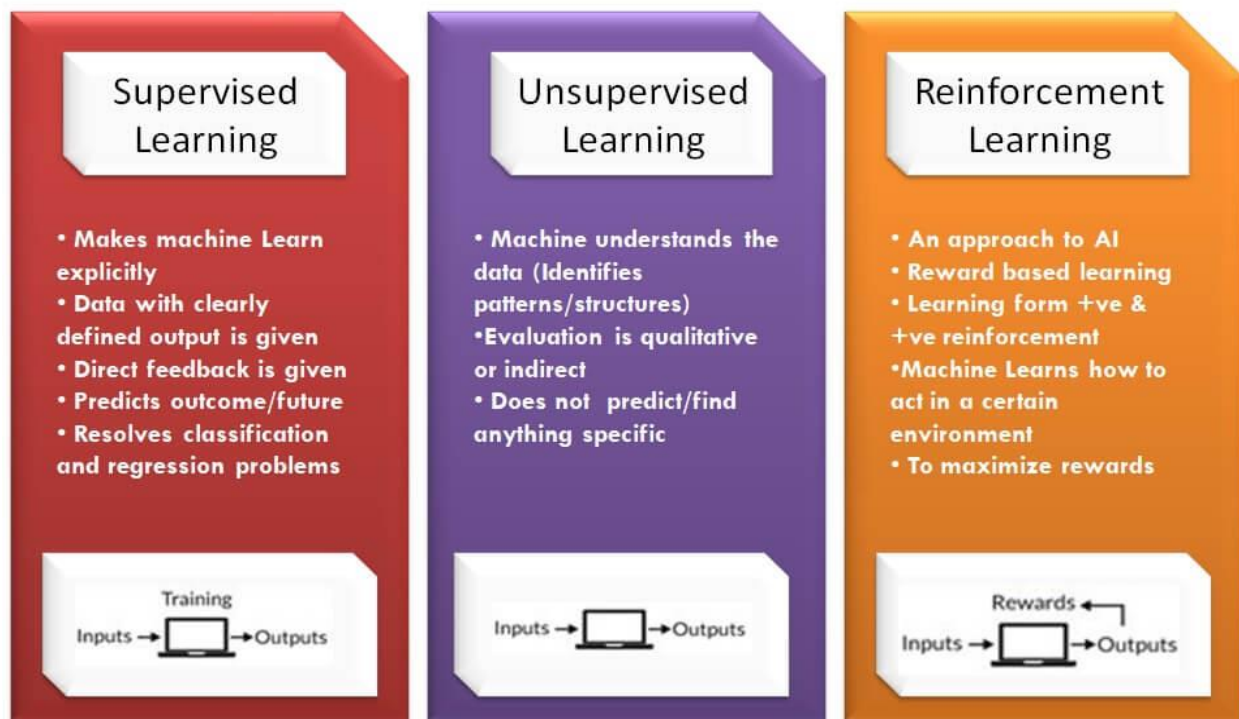
Python offers concise and readable code. While complex algorithms and versatile workflows stand behind machine learning and AI, Python's simplicity allows developers to write reliable systems. Developers get to put all their effort into solving an ML problem instead of focusing on the technical nuances of the language.

NLP means **Natural Language Processing**. Powerfully, deep learning-based NLP allows marketers to gain a detailed understanding of users, generating more valuable insights. Not only does this enable you to widen the group of people you can target, you can also target them with more relevant offers.

9. Supervised Learning with examples.

Supervised learning, also known as supervised machine learning, is a subcategory of machine learning and artificial intelligence.

Types of Machine Learning – At a Glance



Example of Supervised Learning:

One practical example of supervised learning problems is predicting house prices. How is this achieved? First, we need data about the houses: square footage, number of rooms, features, whether a house has a garden or not, and so on. We then need to know the prices of these houses, i.e. the corresponding labels.

Machine learning models cheat sheet

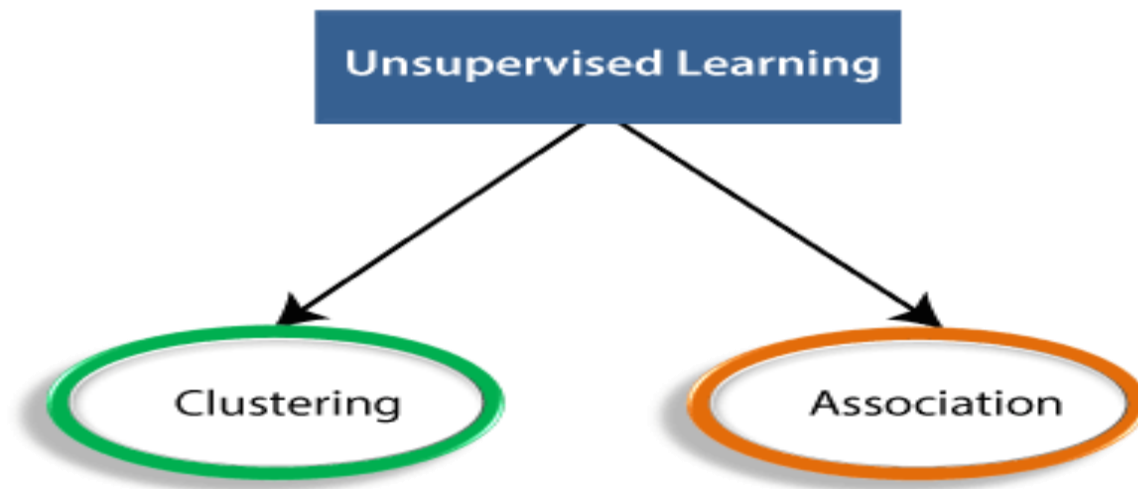
Supervised learning	Unsupervised learning	Semi-supervised learning	Reinforcement learning
<p>Data scientists provide input, output and feedback to build model (as the definition)</p> <p>EXAMPLE ALGORITHMS:</p> <p>Linear regressions</p> <ul style="list-style-type: none"> ■ sales forecasting ■ risk assessment <p>Support vector machines</p> <ul style="list-style-type: none"> ■ image classification ■ financial performance comparison <p>Decision tree</p> <ul style="list-style-type: none"> ■ predictive analytics ■ pricing 	<p>Use deep learning to arrive at conclusions and patterns through unlabeled training data.</p> <p>EXAMPLE ALGORITHMS:</p> <p>Apriori</p> <ul style="list-style-type: none"> ■ sales functions ■ word associations ■ searcher <p>K-means clustering</p> <ul style="list-style-type: none"> ■ performance monitoring ■ searcher intent 	<p>Builds a model through a mix of labeled and unlabeled data, a set of categories, suggestions and example labels.</p> <p>EXAMPLE ALGORITHMS:</p> <p>Generative adversarial networks</p> <ul style="list-style-type: none"> ■ audio and video manipulation ■ data creation <p>Self-trained Naïve Bayes classifier</p> <ul style="list-style-type: none"> ■ natural language processing 	<p>Self-interpreting but based on a system of rewards and punishments learned through trial and error, seeking maximum reward.</p> <p>EXAMPLE ALGORITHMS:</p> <p>Q-learning</p> <ul style="list-style-type: none"> ■ policy creation ■ consumption reduction <p>Model-based value estimation</p> <ul style="list-style-type: none"> ■ linear tasks ■ estimating parameters

10. Unsupervised learning with examples.

Unsupervised learning is helpful for data science teams that don't know what they're looking for in data.

The goal of unsupervised learning is to find the underlying structure of dataset, group that data according to similarities, and represent that dataset in a compressed format.

Example: Suppose the unsupervised learning algorithm is given an input dataset containing images of different types of cats and dogs.



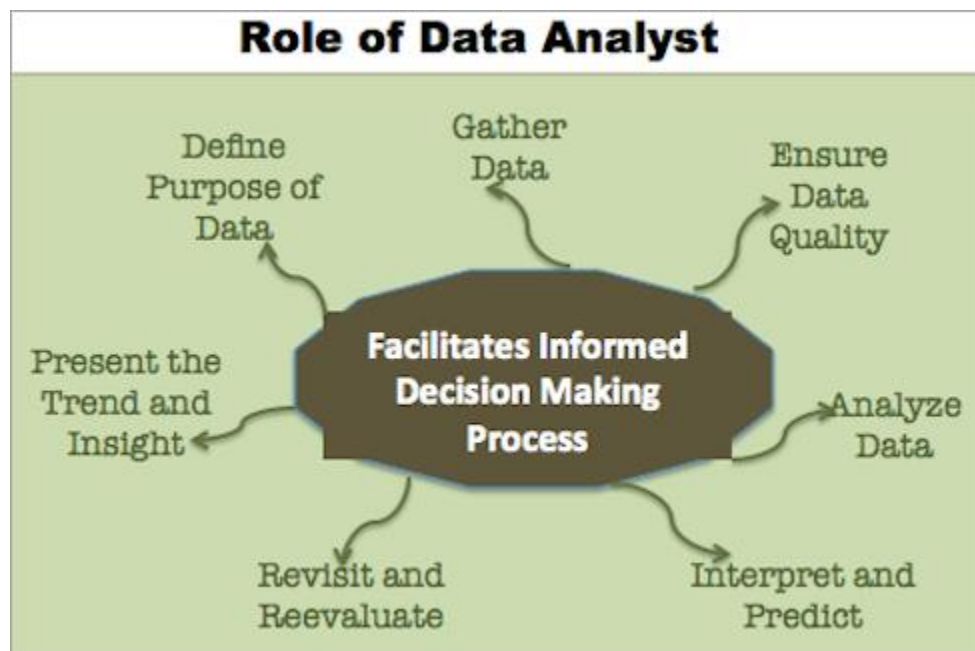
Jobs:

11. Role and responsibilities of a Database Engineer.

- Administer and maintain database security, integrity, and optimization.
- Develop and run periodic and requested reports.
- Develop new databases and applications.
- Identify and resolve performance and capacity issues.
- Establish and maintain backup and recovery processes.
- Maintain associated non-referential data.

12. Role and responsibilities of a Data Analyst.

Data analysts are responsible for analyzing data using statistical techniques, implementing and maintaining databases, gathering data from primary and secondary sources, identifying, analyzing and interpreting trends from the data.



- Medical and Health Care Analyst.
- Market Research Analyst.
- Business Analyst.
- Business Intelligence Analyst



13. Role and responsibilities of a Data Engineer.

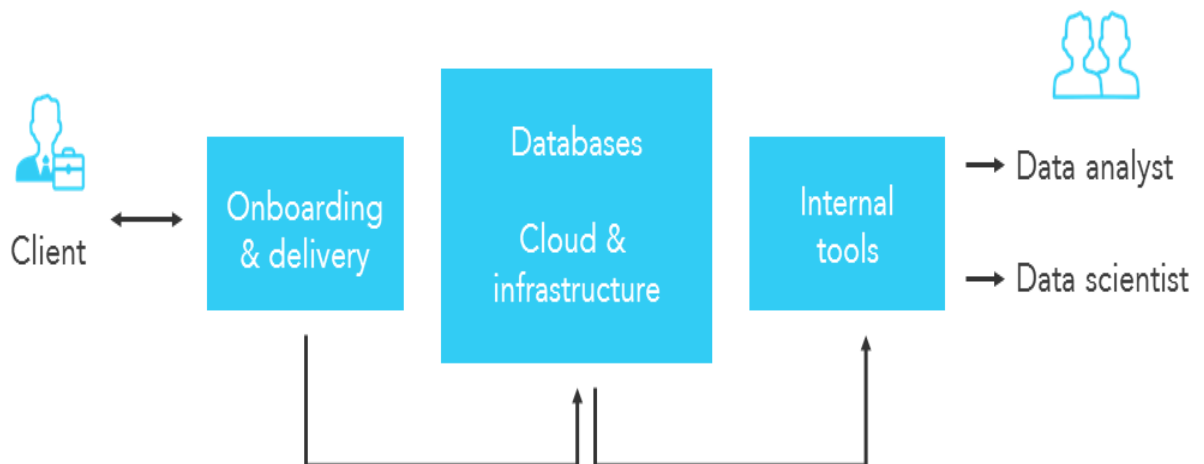
- Work on Data Architecture.
- Collect Data.
- Conduct Research.
- Create Models and Identify Patterns

Responsibilities

- Analyze and organize raw data.
- Build data systems and pipelines.
- Evaluate business needs and objectives.
- Interpret trends and patterns.

- Conduct complex data analysis and report on results.
- Prepare data for prescriptive and predictive modeling.
- Build algorithms and prototypes.

Data engineers' responsibilities



14. Role and responsibilities of a Data Scientist.

Data Engineer	Analytics Engineer	Data Analyst
<ul style="list-style-type: none">• Build custom data integrations• Manage overall pipeline orchestration• Develop & deploy machine learning endpoints• Build and maintain the data platform• Data warehouse performance optimizations	<ul style="list-style-type: none">• Provide clean, transformed data ready for analysis• Apply software engineering best practices to analytics code (ex: version control, testing, continuous integration)• Maintain data documentation & definitions• Train business users on how to use data visualization tools	<ul style="list-style-type: none">• Deep insights work (ex: why did churn spike last month? what are the best acquisition channels?)• Work with business users to understand data requirements• Build critical dashboards• Forecasting

Data scientist responsibilities

- Identifying relevant data sources for business needs.
- Collecting structured and unstructured data.
- Sourcing missing data.
- Organising data in to usable formats.
- Building predictive models.
- Building machine learning algorithms.

- Enhancing the data collection process.

15. Role and responsibilities of a Machine Learning Engineer.

Roles and responsibilities of a machine learning engineer

- Designing ML systems.
- Researching and implementing ML algorithms and tools.
- Selecting appropriate data sets.
- Picking appropriate data representation methods.
- Identifying differences in data distribution that affects model performance.
- Verifying data quality.

Machine Learning Engineer Knowledge and Skills

